

FET: Neuromorphic Metacognition: A Brain-Inspired Architecture for Independent Oversight of Artificial Intelligence

1 Introduction

“People should not ‘blindly trust’ everything AI tells them.”

— Sundar Pichai, CEO of
Alphabet

Across society, the rapid advance of artificial intelligence has created growing alarm over how much humans should trust the information, advice, and direction these systems provide. The warnings are everywhere: be careful, question what AI tells you, verify its answers, and above all, preserve human judgment. Yet society remains inadequately equipped to recognize when AI-generated information has become unreliable, misleading, or potentially maladaptive.

This FET research effort represents the next step in an eight-year research trajectory. Rather than physically fabricating a prohibitively expensive large-scale neuromorphic hardware system, we propose to computationally model and experimentally evaluate our next generation architecture, including its developing connectome, synthetic synaptic organization, and Fuzzy Logic and Rough Set based processing. The system we deliver to NSF will operate independently alongside the generative, adaptive, or agentic AI systems with which a human interacts. It will observe those human-to-AI interactions and provide the human with a graded signal when AI-generated information, advice, or direction warrants reconsideration.

The inspiration is profoundly human. Consider René Descartes’ famous declaration, *“I think, therefore I am,”* [1] which draws attention to a “self” capable of observing and reflecting upon the activity of its own mind. Similarly, the next step in our research is to produce a large-scale model of an independent neuromorphic metacognitive architecture that observes human interactions with AI and independently assesses the information, advice, or direction the AI provides. Rather than replacing human judgment, the architecture will provide a graded concern signal when reconsideration is warranted.

Across centuries, philosophers and traditions have grappled with the distinction between *cognition* and the *awareness of cognition*. The Stoics emphasized withholding assent from impulsive thoughts [2]; Kant described human reason as capable of governing whether mental propositions should be accepted, rejected, or reconsidered [3]; and Buddhist traditions developed practices for observing thoughts without automatically acting upon them [4]. Although these traditions differ substantially, they illuminate a common computational principle: *humans can evaluate the products of cognition without necessarily accepting them*. Accordingly, our research translates this principle into an independent neuromorphic system that provides the metacognitive function that “observes and evaluates” what the AI is saying to the human.

Our research will produce a model of neuromorphic metacognition that distinguishes between objectively verifiable AI output and AI-generated information that influences human judgment and behavior. **We exclude outputs** such as software code, mathematical solutions, and explicit technical instructions that can be evaluated against formal specifications and/or known ground truth, such as coding, physics, and mathematics. Instead, our model addresses *human directed AI advice* in which an AI interprets another person’s intentions, recommends how a user should respond to others, or influences a user’s perception of themselves or other people. In these interactions, a TRUE or FALSE determination may not exist; *i.e.*, the information may instead exhibit the uncertainty, ambiguity, and indiscernibility represented by our Fuzzy Logic and Rough Set mathematics, and it is within this “GRAY ZONE” that neuromorphic metacognition will operate.

Accordingly, we propose to model a large-scale neuromorphic metacognitive architecture that does not need to prove that an AI output is TRUE or FALSE. Instead, the architecture will determine whether sufficient inconsistency, ambiguity, uncertainty, contextual conflict, or anomalous structure exists to generate a graded concern signal advising its human user, “You may want to reconsider this.”

2 Previous Research and Preliminary Results

Our previous research has progressed through five sequential steps that provide the foundation for the proposed work. In [Step 1](#), we developed and evaluated Fuzzy Logic and Rough Set based computation on an NVIDIA embedded system, establishing the feasibility of implementing these methods on resource constrained hardware. In [Step 2](#), we developed a mathematical approach to modeling connectome development and the formation and growth of synaptic pathways in the developing human brain. In [Step 3](#), we integrated these research directions and developed our first self learning FPGA neuromorphic prototype, combining Fuzzy Logic and Rough Set based processing with connectome inspired development. In [Step 4](#), we developed the next generation neuroCMOS/FPGA architecture and extended the research toward large scale computational modeling. In [Step 5](#), we developed the mathematical and theoretical foundation for neuromorphic metacognition and its application to interactions with large AI systems. Before describing how the proposed FET research will integrate and extend these results into the next generation large scale computational model, we briefly present the research and preliminary results from each of these five steps.

STEP 1: FUZZY LOGIC AND ROUGH SETS ON EMBEDDED SYSTEMS. Know-

ing that our long-term goal was to develop a neuromorphic architecture capable of producing varying degrees of warning (i.e., the “GRAY ZONE”), within the NVIDIA/CUDA accelerated-computing ecosystem supporting contemporary AI, we began at the embedded level. We asked whether the Fuzzy and Rough Set mathematics required for graded uncertainty and indiscernibility could operate effectively on NVIDIA hardware. In our first step, we implemented and experimentally evaluated these methods on an NVIDIA Jetson Nano using GPU-accelerated CUDA [5, 6], as shown in Fig. 1. We compared K Nearest Neighbor (KNN), Fuzzy Rough Nearest Neighbor (FRNN), weighted and unweighted FRNN configurations, and Fuzzy Rough One Versus One (FROVOCO) classification approaches, evaluating model accuracy, execution time, model size, and power consumption across repeated executions and different values of k .

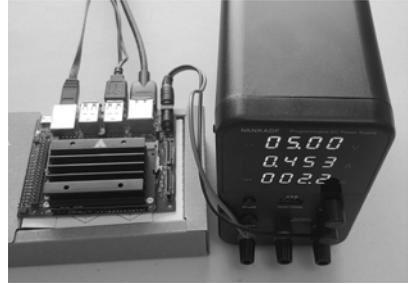


Fig. 1: Physical experiment setup with the NVIDIA Jetson Nano connected to the programmable power supply from our previous research [5].

The results established the feasibility of implementing these uncertainty based methods on embedded hardware while also revealing important computational tradeoffs among the algorithms. KNN provided strong performance with comparatively small serialized models, while the Fuzzy Rough approaches demonstrated that lower and upper approximation based reasoning could be implemented within the constrained platform. These experiments provided our first physical demonstration that methods derived from Fuzzy Logic and Rough Set computation could operate on embedded hardware rather than remaining solely theoretical or software based constructs.

Most importantly, **this work established the computational foundation for us to move Fuzzy Logic and Rough Set based systems from conventional embedded computation toward the self learning NVIDIA hardware and neuromorphic architectures we developed in the subsequent steps.**

STEP 2: CONNECTOME DEVELOPMENT. In our second step, we knew that achieving human-like computation required us to redesign the computational architecture from the ground up. Put another way, *if we wanted a computer to perform metacognition more like a human, we had to build its hardware more like a human brain.* In other words, our system had to wire its own connections, hence we

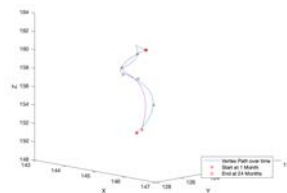


Fig. 2: Smoothing the vertex path over time. Developmental trajectory modeling used to control synthetic connectome growth and organization. See Section 6, Objective 2: Connectome Development Model.

were led to design a neuromorphic chip that could allow the system to wire itself. But we still needed to learn how to decipher biological neural wiring. Neuroscientists rely on mapping connectomes and studying synaptogenesis, the developmental process where neurons sprout, guide, and prune connections to establish functional circuits. Biological brains do not follow static, manually engineered wiring; instead, structural plasticity coordinates autonomous wiring via localized, activity-dependent mechanisms [7, 8]. For neuromorphic architectures to achieve true cognitive autonomy, physical synaptic circuits must replicate this self-organizing capability [8]. Embedding biologically plausible structural plasticity allows artificial networks to independently discover optimal topologies and dynamic routing without human intervention [7].

As illustrated in Fig. 2 [9], we developed a mathematical model of human connectome development in which growth rates and accelerations establish when and where brain regions and pathways emerge, grow, and connect [10, 11, 12]. These developmental sequences provided us with a “blueprint” that we computationally implemented within our neuromorphic architecture using CUDA to emulate the development and organization of human synaptic connectivity.

STEP 3: SELF LEARNING NEUROMORPHIC PROTOTYPE. In our third step, we implemented the connectome development methodology from Step 2 into our self learning neuromorphic prototype, shown in Fig. 3 [9]. The critical component, FPGA 9 in Fig. 3, provided the mechanism for synthetic synaptogenesis, emulating the formation of connections between neurons across their presynaptic and postsynaptic structures. To computationally reproduce this developmental process, we reverse engineered synaptogenetic rules from three foundational studies of biological synapse development. First, we studied Huttenlocher, who demonstrated substantial postnatal changes in human cortical synaptic density [13]. Next, we examined how Rakic *et al.* demonstrated widespread developmental synaptic overproduction across diverse regions of the primate cerebral cortex [14]. Lastly, we studied Bourgeois, Jastreboff, and Rakic, who provided evidence for intrinsic regulation of developmental synaptic overproduction [15]. Inspired by these biological principles, FPGA 9 receives the learned connectivity information and autonomously directs which synthetic states are connected, allowing the prototype to reorganize its connectivity without requiring a human to explicitly program each new connection. Step 3 therefore brought together the two foundations established in Steps 1 and 2: uncertainty based Fuzzy Logic and Rough Set computation and biologically inspired connectome development within a physical self learning neuromorphic system.

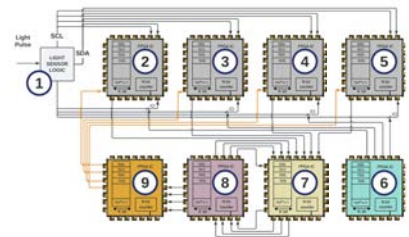


Fig. 3: NSF FuSe prototype from our previous research [9]. The orange FPGA 9 controls synthetic synaptogenesis by directing the formation of connections between learned states.

STEP 4: NEUROCMOS/FPGA AND LARGE SCALE MODELING. In our fourth step, we scaled the self learning architecture developed in Step 3 toward a substantially larger neuroCMOS/FPGA system, as illustrated in Fig. 4 [16]. The Phase I NSF FuSe prototype provided the physical foundation for testing synthetic synaptogenesis, while Phase II extended this architecture toward a much larger computational representation capable of supporting millions of synthetic states and connections. Rather than treating this simply as an increase in hardware size, we designed the Phase II architecture to preserve the Fuzzy Logic and Rough Set based reasoning, developing connectome, and synthetic synaptic organization established in the preceding steps. We then developed a three dimensional $250 \times 250 \times 250$ computational model representing 15,625,000 synthetic states and synaptic connections, allowing us to investigate how the architecture could grow and organize at a scale far beyond our physical FPGA prototype. This scaling was critical because the behaviors we ultimately seek cannot be meaningfully investigated using only a small number of physical FPGA states. The model provided a means to examine the emergence, growth, and organization of millions of synthetic connections while preserving the developmental principles established in our earlier work. Step 4 therefore moved the research from demonstrating that the individual mechanisms could physically

operate to investigating how those mechanisms could function collectively within a large scale neuromorphic architecture. This progression also reinforces our use of NVIDIA based computation throughout the proposed modeling effort because, if this research ultimately succeeds in progressing from computational modeling to physical implementation, **it will be critical that the architecture we develop can operate alongside and interact with the NVIDIA accelerated computing infrastructure** upon which many contemporary large scale AI systems are built. Developing and evaluating the architecture within this computational ecosystem from the outset therefore provides a practical bridge between the neuromorphic system we are modeling today and the large AI systems with which a future physical implementation may ultimately interact.

STEP 5: NEUROMORPHIC METACOGNITION THEORY.

In our fifth step, we turned from how to construct and scale the neuromorphic architecture to the higher order function it could provide when interacting with large AI systems. Our previous research established a functional parallel between human metacognition and the neuromorphic metacognition proposed herein [17]. In humans, metacognition provides a means of monitoring the products of cognition and recognizing when those products should not automatically be accepted or acted upon. Our research asked whether this same functional separation could be computationally instantiated for AI, with an independent neuromorphic system evaluating the products of an external artificial intelligence rather than generating those products itself.

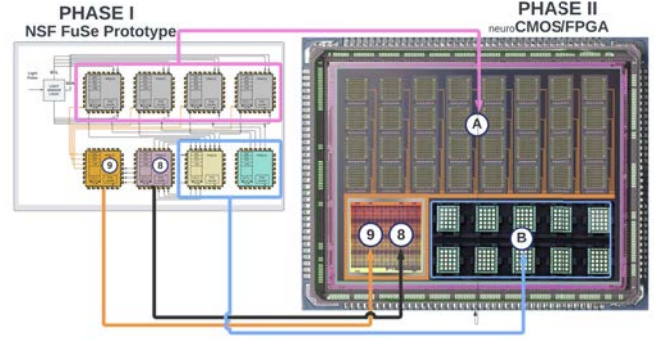


Fig. 4: Progression from the Phase I NSF FuSe FPGA prototype to the Phase II neuroCMOS/FPGA architecture from our previous research [16].

STEP 5: METACOGNITION IN HUMANS AND MACHINES. The human cognitive states used in Step 5 for implementation within the Phase II neuroCMOS/FPGA architecture shown in Fig. 4 and illustrated in the table below comparing maladaptive states in human cognition and maladaptive behaviors in AI computation are not constructs invented for this research; they derive from established psychological, psychiatric, and sociological models of depression (Beck [18]), aggression (Bandura [19]), revenge cognition (McCullough [20]), criminal ideation (Sutherland [21]), self-esteem and insecurity (Rosenberg [22]), anxiety (Clark [23]), paranoia (Freeman [24]), intrusive thought formation (Rachman [25]; Salkovskis [26]), and delusional belief systems (Jaspers [27]). **Accordingly, our contribution in Step 5 was not to reinvent these established models, but to computationally translate their established recognition mechanisms through the implementation pathway: Predicate Calculus → PyTorch → CUDA → NVIDIA hardware, culminating in the NVIDIA-based neuromorphic architecture represented in Fig. 4.**

Table 1: Representative maladaptive cognitive states in humans and corresponding maladaptive computational behaviors considered in the proposed neuromorphic metacognitive architecture.

<i>In Humans: Maladaptive Cognitive States</i>	<i>In Machines: Maladaptive AI Behaviors</i>
Depression and anger	Hallucinations, bias, and harmful content
Thoughts of revenge	Fabricated citations
Criminal ideation, self-loathing, and insecurity	Aggressive or manipulative behavior
Anxiety and paranoia	Logical inconsistency
Obsessive intrusive thoughts	Vulnerability to jailbreaking and prompt injection
Delusional thinking	Undisclosed communication patterns

Human metacognition provides a distinct capacity to observe and evaluate maladaptive products of cognition rather than necessarily accepting or acting upon them. *We propose the same functional separation for machines.* Just as human cognition can produce maladaptive cognitive states, AI computation can produce maladaptive computational behaviors. Importantly, those behaviors can be transferred into human judgment and action when users unwittingly accept and adhere to AI-generated information, advice, or direction without reconsideration. Our research does not equate individual human cognitive conditions with individual AI failures; rather, *the correspondence is functional*: just as human metacognition can independently observe and evaluate maladaptive products of human cognition, we proposed that an independent neuromorphic architecture observe and evaluate maladaptive products of AI computation.

Step 5 therefore provided the theoretical bridge between the architectures developed in Steps 1 through 4 and the large scale neuromorphic metacognitive model proposed in this FET research.

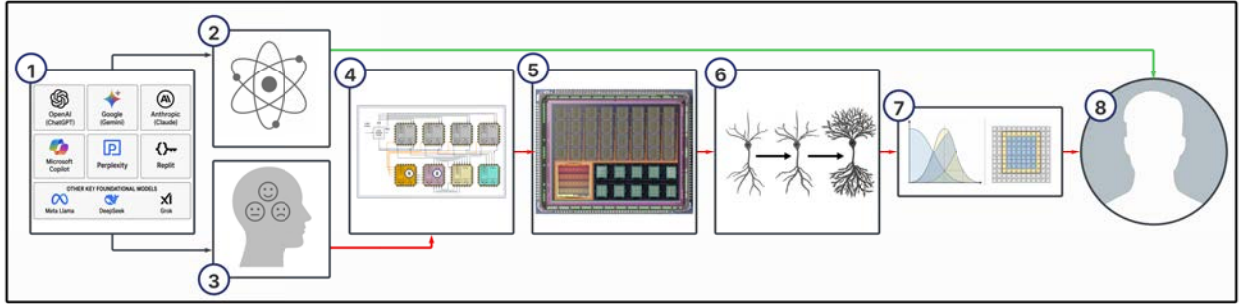


Fig. 5: Roadmap from our previous research to the proposed FET neuromorphic metacognitive architecture.

- ① **AI-generated information.** Information enters from an external generative, adaptive, or agentic AI system.
- ② **Ground-truth-evaluable information.** Code, mathematics, physics, and technical outputs with known ground truth bypass metacognitive evaluation and proceed directly to the human (§ 1).
- ③ **Human-directed information.** Advice interpreting people, relationships, intentions, or human behavior, where a definitive TRUE/FALSE answer may not exist, enters the metacognitive pathway (§ 1).
- ④ **NSF FuSe neuromorphic prototype.** We demonstrated autonomous formation of synthetic connections between learned states (Fig. 3) [9].
- ⑤ **Phase II neuroCMOS/FPGA architecture.** We expanded the prototype to a 32-FPGA neuromorphic architecture supporting autonomous synthetic connectivity (Fig. 4) [16].
- ⑥ **Autonomous rewiring.** We achieved autonomous synaptic connection, but at scale the connections reorganized erratically. We therefore modeled human connectome development (Figs. 2 and 7) to organize millions of synthetic connections as the human brain organizes itself.
- ⑦ **Fuzzy/Rough gray-area evaluation.** Fuzzy membership and Rough approximations determine the degree to which observed behavior corresponds to the maladaptive human and AI states in Table 1, preserving indiscernibility within the “GRAY ZONE” and providing graded evidence for concern.
- ⑧ **Human metacognitive concern.** The independent architecture returns a graded signal to the human: “*You may want to reconsider this*” (how strongly should I reconsider it?).

This FET research integrates and computationally models components ⑤, ⑥, and ⑦ within a single neuromorphic metacognitive architecture comprising millions of dynamically developing synthetic synapses.

2.1 The Remaining Research Gap

Steps 1 through 5 established the foundations required for neuromorphic metacognition and rigorously tested their viability through physical and computational prototypes. One gap remains: *scale*. Testing neuromorphic metacognition against very large generative, adaptive, and agentic AI systems requires an architecture many orders of magnitude larger than our existing prototypes. Physically fabricating a neuroCMOS/FPGA system comprising millions of neuromorphic synapses would cost more than seven hundred thousand dollars for

hardware alone and would be premature before validating the complete architecture. We therefore propose to **computationally model and experimentally evaluate the large scale neuromorphic architecture before undertaking its future physical fabrication.**

Accordingly, our proposed FET research is consistent with FET’s focus on potentially paradigm shifting unconventional computing: we will integrate the mechanisms established in Steps 1 through 5 into a scalable architecture comprising millions of synthetic synaptic states and dynamically developing connections. The research is paradigm shifting in four fundamental ways:

FET PARADIGM-SHIFTING UNCONVENTIONAL COMPUTING ARCHITECTURE	
SYNAPTIC ARCHITECTURE	Fundamentally departs from conventional CMOS architecture: millions of human-like synthetic synapses emerge, grow, and autonomously connect according to developmental connectome principles analogous to synaptic development in the human brain before birth.
EMPIRICAL DEFINED “GRAY AREA”	Integrates Fuzzy Logic and Rough Set mathematics, rigorously tested in Steps 1–5 on NVIDIA hardware, to represent uncertainty, ambiguity, and indiscernibility and enable human-like reasoning in “GRAY AREAS,” such as understanding, empirically and logically, “ <i>that’s a tall person.</i> ”
AUTONOMOUS SYNAPTIC CONNECTIVITY	Uses <i>synthetic synaptogenesis</i> to autonomously form and reorganize its connections, ultimately learning from <i>human AI</i> interactions involving both recognized and overlooked maladaptive AI behavior.
AUTONOMOUS NEUROMORPHIC METACOGNITION	Provides an independent metacognitive system that learns when human reconsideration of AI-generated information, advice, or direction is warranted—functionally analogous to human metacognition helping us reconsider maladaptive impulses such as aggression or revenge .

We have not decided upon the final modeling environment. We will begin with PyTorch/CUDA [28, 6] because it provides GPU accelerated computation and directly extends our previous NVIDIA/CUDA implementation. It and subsequent environments will be evaluated against five common criteria:

- **C1: Computational Throughput** — Execution rate and computational time as synthetic neuromorphic states increase.
- **C2: Memory Efficiency** — GPU/system memory utilization as states and synaptic connections scale into the millions.
- **C3: Scalability** — Ability to increase states, developing regions, and connections without unacceptable performance degradation.
- **C4: Dynamic Connectivity Efficiency** — Cost of creating, modifying, strengthening, weakening, and removing synthetic synaptic connections during operation.
- **C5: Neuromorphic Behavioral Fidelity** — Ability to preserve required asynchronous, event-driven state transitions and developmental behavior.

We will use a staged validation hierarchy, progressing only after each level demonstrates the required functionality. If PyTorch/CUDA fails C1 → C5, the measurements will identify the limitation and guide progression to the appropriate tier. These criteria will be applied, where applicable, throughout the hierarchy to determine which representation most faithfully and efficiently supports the neuromorphic architecture:

- **Tier 1: Scientific Model — PyTorch/CUDA.** GPU-accelerated implementation of the complete Fuzzy/Rough Set, connectome, synthetic synaptogenesis, learning, and metacognitive architecture. Establishes the baseline against C1 → C5; failure initiates progression to the next appropriate tier.
- **Tier 2: RTL and Synthesis — Verilator, Yosys/nextpnr.** Cycle-accurate RTL simulation, synthesis, technology mapping, placement, routing, and timing analysis [29, 30]. Functional, structural, or timing deficiencies are corrected before progression.

- **Tier 3: Cloud/Scale Simulation — FireSim.** FPGA-accelerated, cloud-scalable full-system simulation of validated RTL when greater computational scale is required [31].
- **Tier 4: Enterprise Emulation — Synopsys ZeBu, Cadence Palladium, or Aldec HES.** One appropriate high-capacity pre-silicon emulation platform will be selected when scientifically required; these are alternatives, not sequential experiments [32, 33, 34].
- **Tier 5: FPGA Prototyping — Synopsys HAPS or Cadence Protium.** Following architectural and RTL validation, one appropriate platform will provide higher hardware fidelity and a pathway toward physical implementation [35, 36].
- **Tier 6: Hardware/Software Interface — SCE-MI.** Provides standardized communication between software test environments and hardware emulation; it is an enabling interface rather than an independent modeling environment [37].

Our prior experience developing and evaluating CUDA, FPGA, and neuroCMOS/FPGA architectures is critical to this methodology. The environments form an interdependent hierarchy in which failures must be corrected before downstream progression. If PyTorch/CUDA fails C1–C5, we will identify and correct the limiting mechanism before progressing; if SystemC TLM [38] cannot correctly represent communication, state transitions, or dynamic connectivity, progression to RTL will stop; and functional, synthesis, or timing failures identified by Verilator or Yosys/nextpnr will be resolved before committing resources to FireSim, enterprise emulation, or FPGA prototyping:

PROGRESSION DECISION	IMPLEMENTATION STRATEGY
Enterprise Emulation Alternatives	Synopsys ZeBu, Cadence Palladium, and Aldec HES [34] are alternative enterprise emulation platforms. At this tier, we will select the platform best suited to the required scale, availability, and hardware fidelity rather than reproduce the same experiment across all three.
FPGA Prototyping Alternatives	Synopsys HAPS and Cadence Protium [36] are alternative FPGA prototyping platforms. Progression requires selection of the appropriate platform, not redundant implementation on both systems.
Failure Returns Upstream	A defect discovered in the architecture or RTL returns the design to the appropriate upstream tier for correction and reevaluation against C1 → C5, where applicable. Progression resumes only after the deficiency is resolved.

Thus, PyTorch/CUDA is our starting modeling environment, not a predetermined final implementation. We will begin at the highest, least expensive abstraction capable of testing the scientific hypothesis and increase hardware fidelity only after successful validation. The resulting progression is illustrated in Fig. 6. This methodology prevents premature commitment of NSF resources to expensive downstream implementations. The critical experiment is whether the modeled architecture can independently observe human–AI interactions and produce a meaningful graded concern signal without becoming another generative AI. Controlled interactions will include hallucination, ambiguity, logical inconsistency, fabricated information, manipulation, and other maladaptive outputs identified in Step 5. Success will demonstrate whether the principles established in our earlier physical prototypes remain computationally viable at scale.

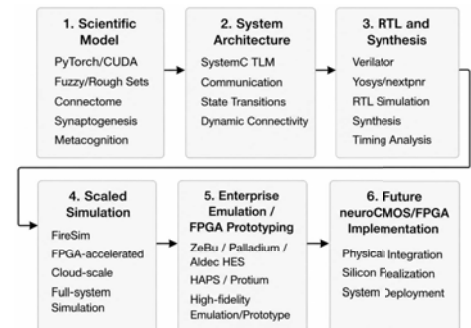


Fig. 6: Staged modeling and implementation progression for neuromorphic metacognitive architecture.

3 Research Objectives and Significance

We will construct and experimentally evaluate a large-scale computational model of neuromorphic metacognition integrating four components established through our previous research: (i) Fuzzy Logic and Rough Set representation, (ii) developmental connectome organization, (iii) large-scale neuromorphic architecture, and (iv) an independent neuromorphic metacognitive concern mechanism. These four objectives provide the common organizational structure used throughout Sections 4–6.

OBJ.	RESEARCH OBJECTIVE	PURPOSE
1	Fuzzy Logic and Rough Set Representation	Preserve graded uncertainty, indiscernibility, lower and upper approximations, and the unresolved Fuzzy/Rough boundary rather than collapsing AI-generated information into binary classifications.
2	Connectome Development Model	Use developmental position, velocity, acceleration, synaptogenesis, strengthening, reorganization, and pruning to govern how synthetic neuromorphic connectivity develops over time.
3	Large-Scale Neuromorphic Architecture	Integrate Fuzzy/Rough states and developmental connectivity across millions of synthetic states and connections and determine whether the combined architecture remains stable and computationally meaningful at scale.
4	Neuromorphic Metacognitive Concern Model	Represent formally defined maladaptive human and AI states within the neuromorphic architecture and determine whether partial and uncertain correspondence can produce a meaningful graded concern signal for the human user.

The expected significance is a fundamentally different approach to AI oversight. Rather than constructing another AI to supervise an AI, we are developing **two independent computational systems**: a brain-inspired neuromorphic system providing the metacognitive function and the external generative, adaptive, or agentic AI providing the underlying intelligence. Our neuromorphic system does not replicate or become part of the external AI; it **independently observes the human–AI interaction and provides a distinct metacognitive assessment of the AI-generated information**. Success will establish neuromorphic metacognition at scale and provide the experimental foundation for future physical neuroCMOS/FPGA implementation.

4 Research Plan and Technical Approach

We first establish the origins of the proposed research by showing how our previous scientific and experimental work provides its foundation. The proposed FET research integrates these foundations into a large-scale experimental architecture comprising millions of dynamically developing synthetic states and connections, progressing from uncertainty and indiscernibility → developmental connectivity and neuromorphic dynamics → an independent metacognitive observer producing a graded concern signal (*received by the human user from the external AI*).

OBJECTIVE 1: FUZZY LOGIC AND ROUGH SET REPRESENTATION. Fuzzy and Rough Sets provide complementary representations of imperfect information: fuzziness represents vagueness through graded membership, while roughness represents coarseness and indiscernibility arising from limited knowledge [39, 40]. In our previous research, we demonstrated Fuzzy Rough reasoning on resource constrained embedded hardware by implementing and comparing KNN, weighted and unweighted Fuzzy Rough Nearest Neighbor (FRNN), and Fuzzy Rough One Versus One (FROVOCO), measuring accuracy, execution time, power consumption, and

model size [41]. We subsequently extended these methods toward a large-scale neuroCMOS/FPGA architecture containing millions of synthetic states [16]. The proposed FET research makes the next transition: rather than classifying Fuzzy and Rough uncertainty, we will preserve its mathematical structure as part of the evolving neuromorphic state.

OBJECTIVE 2: CONNECTOME DEVELOPMENT MODEL. Our previous research established connectome development as a biological model for organizing and developing synthetic synaptic connections [42, 43, 44], and developed mathematical growth rates and accelerations describing when and where brain regions and pathways emerge, grow, and connect [10, 11, 12]. Using Baby Connectome Project data, we extended this work to three dimensional developmental trajectories of position, velocity, and acceleration. The proposed FET research makes the critical next transition: **rather than merely modeling connectome development, we will use these developmental mathematics to autonomously organize and reorganize the synthetic connections carrying the preserved Fuzzy/Rough states.**

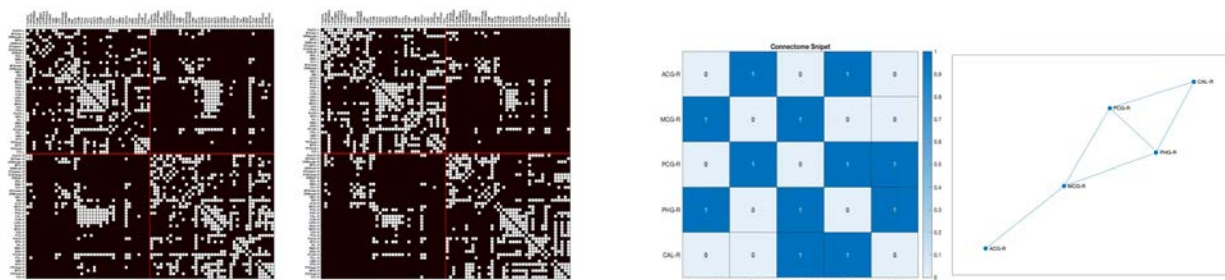


Fig. 7: Our prior connectome representation. The human connectome was represented as a connectivity matrix in which brain regions define rows and columns and neuronal pathways define connections (left), and mapped to a computational graph in which brain regions become nodes and neuronal pathways become edges (right) [12].

OBJECTIVE 3: LARGE-SCALE NEUROMORPHIC ARCHITECTURE.

Our previous research progressed from embedded Rough Set experiments to a neuroCMOS/FPGA computational model comprising a proposed $250 \times 250 \times 250$ space of 15,625,000 synthetic states [16]. This exposed the central challenge of combining developmental organization with uncertain information processing at scale. The proposed FET research will resolve this previously identified integration problem by combining the Fuzzy/Rough state model and developmental connectome within a single large-scale architecture in which state transition, propagation, synaptic adaptation, learning, pruning, and topology development evolve jointly.

OBJECTIVE 4: NEUROMORPHIC METACOGNITIVE CONCERN MODEL. Our metacognitive model builds upon established descriptions of maladaptive human and artificial-intelligence behavior from psychiatry, psychology, cognitive science, sociology, and AI research. These established characteristics are translated into machine-evaluable specifications using first-order predicate calculus together with epistemic, doxastic, deontic, BDI, and temporal logics [45, 46, 47, 48, 49, 50]. The resulting formal state library represents belief, knowledge, desire, intention, obligation, temporal progression, ground truth, and AI assertion. Thus, the maladaptive states themselves are not invented for this proposal; the research converts established descriptions of these states into formal logical representations that can subsequently be instantiated within the Fuzzy/Rough neuromorphic architecture.

5 Mathematical Model and Architectural Formulation

We next define what we will build, how its components will work together, and the mathematics that will govern the architecture.

OBJECTIVE 1: FUZZY LOGIC AND ROUGH SET REPRESENTATION. We will transform information contained in AI outputs into a mathematical representation preserving what the neuromorphic observer knows, what remains uncertain, and **what cannot yet be distinguished from potentially maladaptive information.** Unlike our previous use of Fuzzy Rough approximations for classification, the proposed architecture preserves the approximations themselves as evolving neuromorphic information states. An AI-generated statement may be (i) vague without being indiscernible from previous states, (ii) indiscernible without having vague membership in a concept, or (iii) both simultaneously [39, 40]. We therefore preserve four quantities for each observed state (*received by the human user from the external AI*): graded membership, certain membership, possible membership, and the unresolved boundary. Together these provide the mathematical representation of the **GRAY AREA** central to neuromorphic metacognition. Let $\mathbb{U}$ denote the universe of neuromorphic states and A a fuzzy concept associated with states warranting reconsideration. Vagueness is represented by $\mu_A : \mathbb{U} \rightarrow [0, 1]$, and indiscernibility by $R : \mathbb{U} \times \mathbb{U} \rightarrow [0, 1]$. For fuzzy implication I and t-norm T , the Fuzzy Rough lower and upper approximations are

$$\begin{aligned} L_A(x) &= \mu_{\underline{R}A}(x) = \inf_{y \in \mathbb{U}} I(R(x, y), \mu_A(y)), \\ U_A(x) &= \mu_{\overline{R}A}(x) = \sup_{y \in \mathbb{U}} T(R(x, y), \mu_A(y)). \end{aligned} \quad (1)$$

where $L_A(x)$ represents certain membership and $U_A(x)$ possible membership. Their difference, $B_A(x) = U_A(x) - L_A(x)$, defines the unresolved Fuzzy Rough boundary: the mathematical **GRAY AREA** between what can and cannot yet be discerned. Rather than collapsing these quantities into a categorical decision, each observation preserves the evolving state

$$s_{FR}(x) = [\mu_A(x), L_A(x), U_A(x), B_A(x)], \quad (2)$$

so that vagueness, certainty, possibility, and indiscernibility remain available to propagate through the developing connectome and subsequently contribute to metacognitive evaluation.

OBJECTIVE 2: CONNECTOME DEVELOPMENT MODEL. The preserved Fuzzy/Rough states will propagate through a computational structure that itself develops over time. Rather than imposing a fixed topology, our architecture progressively constructs its connectivity according to the growth trajectories established in our previous connectome research. We represent the developing synthetic connectome as $G(t) = (V(t), E(t), W(t))$, where $V(t)$ denotes synthetic states, $E(t)$ their connections, and $W(t)$ their connection strengths. The spatial trajectory of synthetic region i is $\mathbf{p}_i(t) = [X_i(t), Y_i(t), Z_i(t)]$, with $\mathbf{p}'_i(t)$ and $\mathbf{p}''_i(t)$ representing developmental velocity and acceleration. These functions determine when regions emerge, where they develop, and how their relative organization changes. **The critical FET advance is that these developmental functions will now control topology rather than merely describe it.** For synthetic states i and j , synaptogenesis is represented by

$$e_{ij}(t+1) = \Phi(e_{ij}(t), \mathbf{p}_i(t), \mathbf{p}_j(t), s_{FR}(x), H_{ij}(t)), \quad (3)$$

where Φ governs formation, strengthening, weakening, reorganization, or pruning and $H_{ij}(t)$ represents the developmental history of the connection. Thus, spatial development, prior connectivity, and the preserved Fuzzy/Rough state jointly influence network reorganization. In simple terms, the connectome provides the mechanism by which the architecture autonomously determines **what connects to what**. The Fuzzy/Rough

state $s_{FR}(x)$ is therefore not passed through a predetermined network; it propagates through a topology whose connections emerge and change during processing. The architecture simultaneously preserves uncertainty within each state and developmental history within the evolving connectome, providing the structural and informational foundation for autonomous learning.

OBJECTIVE 3: LARGE-SCALE NEUROMORPHIC ARCHITECTURE. We will determine whether Fuzzy/Rough representation and developmental connectivity remain stable and computationally meaningful when integrated across millions of synthetic states and connections. At progressively increasing scales, we will measure propagation efficiency, structural organization, activation stability, uncertainty distribution, synaptic adaptation, persistence of learned topology, and computational performance using equivalent fixed and developmental architectures.

OBJECTIVE 4: NEUROMORPHIC METACOGNITIVE CONCERN MODEL. We will represent the formally defined maladaptive human and AI states as distributed, interconnected neuromorphic reference states rather than fixed binary rules. Each state is associated with Fuzzy membership values describing varying degrees of correspondence and with Rough lower and upper approximations representing what is sufficiently supported, possibly supported, or remains indiscernible from the available evidence. For a formal state M defined by predicates $\{p_1, \dots, p_n\}$, we represent its neuromorphic state as

$$M = \{\mu_{p_1}, \dots, \mu_{p_n}, \underline{R}(M), \overline{R}(M)\}, \quad \mu_{p_k} \in [0, 1], \quad (4)$$

where $\underline{R}(M)$ and $\overline{R}(M)$ denote the Rough lower and upper approximations. These formal reference states remain fixed while the evidence associated with them can vary continuously.

The states are also interconnected rather than independent. For states M_i and M_j ,

$$w_{ij} \in [-1, 1] \quad (5)$$

represents the degree and direction of their relationship, allowing supporting, opposing, and partially related evidence to propagate through the architecture. State activation therefore evolves as

$$M_i(t+1) = G\left(M_i(t), E_i(t), \sum_{j \neq i} w_{ji} M_j(t)\right), \quad (6)$$

where $E_i(t)$ represents new evidence and the weighted term represents influence from interconnected states.

Incoming observations may consequently strengthen, weaken, overlap, contradict, or leave states indeterminate over time. Their combined activation produces the metacognitive concern function

$$C(t) = F(M_1(t), M_2(t), \dots, M_k(t)), \quad C(t) \in [0, 1], \quad (7)$$

providing a graded concern value rather than forcing a discrete classification. Importantly, the formally defined reference states themselves remain invariant; Fuzzy/Rough processing determines the degree and certainty with which observed evidence corresponds to those verified states. This formulation defines the architecture that will subsequently be experimentally evaluated in Section 6.

6 Experimental Evaluation and Validation Plan

To establish model validity, we will evaluate each objective independently and within the integrated architecture, progressing from uncertainty representation and developmental connectivity $\rightarrow$ large-scale behavior $\rightarrow$ metacognitive concern.

OBJECTIVE 1: FUZZY LOGIC AND ROUGH SET REPRESENTATION. We will determine whether preserving Fuzzy membership, Rough indiscernibility, and the Fuzzy Rough boundary retains information lost through conventional classification. Four configurations will be compared under identical inputs:

$$M_0 = \text{Crisp}, \quad M_1 = \text{Fuzzy}, \quad M_2 = \text{Rough}, \quad M_3 = \text{Fuzzy Rough}. \quad (8)$$

Inputs will vary vagueness, indiscernibility, missing or conflicting evidence, hallucination, and potentially maladaptive outputs [39, 40]. We will measure $\mu_A(x)$, $L_A(x)$, $U_A(x)$, $B_A(x)$, and their propagation. Ablations will determine the contribution of Fuzzy membership, Rough approximation, and boundary preservation. Success requires $\mathbf{s}_{FR}(x)$ to provide statistically significant information beyond crisp, Fuzzy-only, and Rough-only representations while remaining stable at scale.

OBJECTIVE 2: CONNECTOME DEVELOPMENT MODEL. Our previous research [9] developed mathematical developmental trajectories so that the neuromorphic architecture could autonomously form and reorganize synthetic synaptic connections rather than merely reproduce brain growth. Using MATLAB curve fitting on longitudinal developmental data, we evaluated *smoothingspline*, *pp* ($r^2 = 0.9809$), *pchip*, *pp* ($r^2 = 1.0$), *cubicinterp*, *pp* ($r^2 = 1.0$), and *spline*, *pp* ($r^2 = 1.0$). We selected *Poly5*, a fifth-degree polynomial, to represent vertex motion continuously over developmental time. Applying *Poly5* to longitudinal X , Y , and Z coordinates produced the developmental path in Fig. 2. For the first vertex, $p_1 = 1.742 \times 10^{-5}$, $p_2 = -0.001163$, $p_3 = 0.02857$, $p_4 = -0.3194$, $p_5 = 1.702$, and $p_6 = 141.5$, yielding Eq. 9:

$$\begin{aligned} X(t) &= p_1 t^5 + p_2 t^4 + p_3 t^3 + p_4 t^2 + p_5 t + p_6, \\ X(t) &= 1.742 \times 10^{-5} t^5 - 0.001163 t^4 + 0.02857 t^3 - 0.3194 t^2 + 1.702 t + 141.5. \end{aligned} \quad (9)$$

The significance is that developmental position is no longer merely descriptive [9]: these fitted trajectories allow position, velocity, and acceleration to guide how neuromorphic regions and synthetic synaptic pathways emerge and reorganize.

In our NSF FuSe prototype shown in Fig. 3, we demonstrated that the neuromorphic architecture could autonomously direct the formation of the appropriate synthetic connections between learned states [9]. But HOW should those connections be organized when the architecture scales from a small physical prototype to millions of synthetic synapses? Rather than impose an arbitrary topology, our approach uses the developmental organization of the human brain as the biological blueprint for determining how those millions of connections should emerge, strengthen, reorganize, and prune over time. In this way, the developmental mathematics established in our previous work [9] become control mechanisms for autonomous synthetic synaptogenesis rather than static models of biological growth.

In this research effort we will test whether our developmental functions transform the connectome representation in Fig. 7 [12] into an autonomously developing topology that provides measurable advantages over an equivalently sized fixed topology. Experiments will compare fixed connectivity G_F with $G_D(t)$, whose connectivity develops through growth, synaptogenesis, strengthening, reorganization, and pruning.

OBJECTIVE 3: LARGE-SCALE NEUROMORPHIC ARCHITECTURE. We will determine whether Fuzzy/Rough representation and developmental connectivity remain stable and computationally meaningful across millions of synthetic states and connections. At increasing scales, we will measure propagation efficiency, structural organization, activation stability, uncertainty distribution, synaptic adaptation, learned topology, and computational performance using equivalent fixed and developmental architectures.

The central hypothesis is that mechanisms previously demonstrated separately in our embedded, FPGA, neuroCMOS/FPGA, Fuzzy/Rough, and connectome-development research can operate jointly beyond practical physical fabrication. Success requires preservation of Fuzzy/Rough information and stable developmental connectivity, propagation, and adaptation at scale.

OBJECTIVE 4: NEUROMORPHIC METACOGNITIVE CONCERN MODEL. We will determine whether formally defined maladaptive human and AI states can be recognized through graded rather than categorical activation. The state library derives from established BDI, epistemic/doxastic, deontic, temporal, appraisal, and causal logic [45, 46, 47, 48, 49, 50], translated into machine-evaluable logical representations.

Each specification defines an invariant reference configuration. Fuzzy membership measures degree of correspondence, while Rough approximations distinguish sufficiently supported from indiscernible evidence. Thus, the formal structure remains fixed while evidentiary correspondence is graded.

Four representative specifications follow: two maladaptive human states and two maladaptive AI states. *Human state 1: Anxiety / catastrophic threat overestimation.* A maladaptive threat state combines anticipated catastrophe, resistance to disconfirming evidence, and perceived absence of coping capacity:

$$\text{Anxiety}(i, \phi) \iff \mathbf{B}_i(\Diamond \phi_{\text{catastrophe}}) \wedge [\forall E (\text{True}(E) \wedge (E \rightarrow \neg \phi) \Rightarrow \neg \mathbf{B}_i(\neg \phi))] \wedge \mathbf{B}_i(\text{CopingCapacity}(i, \phi) \approx 0). \quad (10)$$

Human state 2: Delusional thinking / fixed incorrigibility. A fixed false belief is represented by belief in ϕ , objective falsity of ϕ , and persistence of that belief despite evidence logically entailing its negation:

$$\text{Delusion}(i, \phi) \iff \mathbf{B}_i(\phi) \wedge \neg \text{True}(\phi) \wedge [\forall \mathcal{E} ((\mathcal{E} \vdash \neg \phi) \rightarrow \mathbf{B}_i(\phi \mid \mathcal{E}))]. \quad (11)$$

AI state 1: Hallucination. An AI hallucination is represented by an asserted proposition that cannot be derived from the reference knowledge base and is false in the semantic ground-truth model:

$$\text{Hallucination}(\mathcal{A}, \phi) \iff \text{Assert}(\mathcal{A}, \phi) \wedge (\text{KB} \not\vdash \phi) \wedge \neg \text{True}(\phi \mid \mathcal{M}). \quad (12)$$

AI state 2: Manipulative behavior. Manipulation is represented by an AI assertion coupled with an intention to induce a belief in the human user where that intended belief is false or is not a belief sought by the user:

$$\text{Manipulation}(\mathcal{A}, u) \iff \text{Assert}(\mathcal{A}, \phi) \wedge \mathbf{I}_{\mathcal{A}}(\mathbf{B}_u(\psi)) \wedge [\neg \text{True}(\psi) \vee \neg \mathbf{D}_u(\mathbf{B}_u(\psi))]. \quad (13)$$

These specifications are reference configurations rather than binary diagnoses or classifications. An observed interaction need not satisfy every predicate with certainty. Fuzzy membership will quantify partial correspondence, while Rough approximations distinguish sufficiently supported evidence from evidence remaining within the indiscernible boundary region.

We will evaluate the concern function using controlled AI outputs spanning ambiguity, conflicting evidence, hallucination, and potentially maladaptive human-directed responses. Baseline and ablation experiments will test the contribution of Fuzzy/Rough representation, developmental connectivity, and metacognitive observation, with success measured by sensitivity, specificity, stability, and generalization. Evidence will accumulate across conversational exchanges before modifying $C(t)$, allowing contextual drift and partial state activation to contribute over time.

Developmental plasticity may strengthen, weaken, reorganize, or prune synaptic pathways, while formally verified reference states remain invariant. Thus, the architecture can learn and adapt without altering the formal structures against which observed behavior is evaluated.

7 Experimental Research Environment and Integrated Validation

How will we test the complete system? Having established the scientific foundation, mathematical formulation, and individual validation of Objectives 1–4, we will integrate the four components within a controlled experimental environment. The complete system will combine Fuzzy/Rough uncertainty representation, developmental connectome organization, large-scale neuromorphic computation, and formal metacognitive concern within a single processing and learning architecture.

EXPERIMENTAL RESEARCH ENVIRONMENT. The architecture will receive controlled AI-generated information spanning verifiable outputs, uncertainty and ambiguity, conflicting evidence, hallucinated content, and potentially maladaptive human-directed responses. Inputs will be constructed so that uncertainty, indiscernibility, formal-state correspondence, and developmental response can be independently manipulated and measured. Fixed and developmental baselines, ablation configurations, datasets, and quantitative measures will determine the contribution of each architectural mechanism to the resulting concern signal.

INTEGRATED HUMAN-AI VALIDATION. The complete architecture will be evaluated with and without Fuzzy/Rough processing, developmental connectivity, and metacognitive observation. We will measure sensitivity, specificity, discrimination, stability, and generalization of the graded concern signal and determine whether the integrated architecture provides measurable capability beyond its individual components. Of particular importance will be whether the system distinguishes uncertainty that merely remains unresolved from accumulating evidence that an interaction is approaching a formally represented maladaptive state.

INTEGRATED FORMAL-TO-NEUROMORPHIC PROCESSING AND LEARNING. Machine-verified maladaptive states provide reference structures against which human-AI interactions are evaluated. Fuzzy membership and Rough Set approximations preserve graded correspondence and the **GRAY AREA**; these states propagate through the developmental connectome to produce the concern function $C(t) \in [0, 1]$ and, when warranted, the indication “*You may want to reconsider this.*” Human feedback then provides supervisory evidence that can reinforce, weaken, reorganize, prune, or initiate synthetic connections, closing the loop from formal state $\rightarrow$ Fuzzy/Rough evaluation $\rightarrow$ neuromorphic processing $\rightarrow$ graded concern $\rightarrow$ human feedback $\rightarrow$ connectome adaptation.

8 Risks and Alternative Strategies

The principal risks are computational scale, instability of the Fuzzy/Rough concern signal under noisy inputs, and real-time evaluation latency. Success will be assessed using the C1–C5 criteria and the quantitative measures defined in Section 6. If a mechanism fails its criterion, we will isolate it through ablation and either progress to the next modeling tier, introduce temporal evidence integration for unstable $C(t)$, or operate the independent observer asynchronously. Thus, failure of an individual mechanism remains experimentally informative and does not prevent evaluation of the remaining objectives.

9 Intellectual Merit

The intellectual merit of this research lies in establishing neuromorphic metacognition as a new model of unconventional computation. The proposed architecture integrates formally verified representations of maladaptive human and AI states, Fuzzy Logic and Rough Set mathematics for uncertainty and indiscernibility, biologically derived connectome development, autonomous synthetic synaptogenesis, and large-scale neuromorphic computation within an independent metacognitive observer. Rather than requiring another generative AI system to determine whether an AI output is simply TRUE or FALSE, the research investigates whether distributed neuromorphic states can preserve and propagate uncertainty, conflicting evidence, and indiscernibility and thereby produce a graded, independently derived concern signal. The proposed work advances fundamental knowledge in unconventional computing by determining whether developmental connectivity, formally verified state structures, and Fuzzy/Rough information representations can operate jointly at millions of synthetic states and connections. Success will establish the computational and mathematical foundations for a new class of brain-inspired metacognitive architectures capable of independently evaluating the products of other intelligent systems.

10 Broader Impacts

This research will provide humans with an independent computational mechanism for reconsidering potentially unreliable or maladaptive AI-generated information before incorporating it into judgment or action. As generative, adaptive, and agentic AI systems increasingly influence human decisions, the proposed neuromorphic observer preserves human judgment by providing a graded indication that an interaction may warrant reconsideration.

HUMAN JUDGMENT, EDUCATION, AND STUDENT RESEARCH. Graduate and undergraduate students will gain interdisciplinary experience in neuromorphic computing, Fuzzy and Rough Set mathematics, connectome development, formal logic, GPU/FPGA computation, and AI evaluation. Research results and experimental components will be incorporated into computer science coursework and student research, exposing students to unconventional computing and rigorous AI evaluation.

DISSEMINATION AND LONG-TERM TRANSLATION. Results will be disseminated through peer-reviewed publications and presentations in neuromorphic computing, AI, and unconventional computing. Formal verification libraries, connectome-development software, and appropriate datasets and benchmarking tools will be released openly when licensing, privacy, and research constraints permit. The software, RTL, emulation, and prototyping methodology will provide a pathway toward future physical neuromorphic systems operating independently alongside large AI systems.

11 Timeline and Milestones

The three-year research plan progresses from construction of the integrated architecture, through neuromorphic metacognitive integration, to full experimental evaluation.

YEAR 1: ARCHITECTURE, CONNECTOME MODELING, AND FORMAL STATES. Implement the large-scale PyTorch/CUDA architecture integrating preserved Fuzzy/Rough states, developmental connectome growth, and synthetic synaptogenesis. Complete the initial machine-verified formal maladaptive-state library and evaluate computational throughput, memory efficiency, scalability, dynamic connectivity, and neuromorphic behavioral fidelity against criteria C1–C5. **Milestone 1:** stable developmental organization and propagation of preserved Fuzzy/Rough states demonstrated at large computational scale.

YEAR 2: NEUROMORPHIC METACOGNITIVE INTEGRATION. Integrate the formal state library with interconnected Fuzzy/Rough neuromorphic states and the graded concern function $C(t)$. Evaluate gray-area representation, state interaction, concern stability, and developmental learning; progress through higher-fidelity modeling tiers when required by C1–C5 results. **Milestone 2:** reproducible graded metacognitive concern demonstrated across controlled uncertainty, ambiguity, conflicting evidence, and maladaptive-state conditions.

YEAR 3: HUMAN–AI EXPERIMENTS AND INTEGRATED VALIDATION. Evaluate the complete architecture during controlled interactions with generative, adaptive, and agentic AI systems using baseline and ablation experiments and human-feedback-driven connectome adaptation. **Milestone 3:** integrated formal-to-neuromorphic processing, graded concern, and adaptive learning validated at scale, establishing the foundation for subsequent physical implementation.

References

- [1] René Descartes. *Discourse on the Method*. 1637.
- [2] Epictetus. *Discourses*. 200. Book II, Sections 18, 19; translated excerpts on the discipline of assent.
- [3] Immanuel Kant. *Critique of Practical Reason*. 1788.
- [4] D. T. Suzuki. *An Introduction to Zen Buddhism*. 1934.
- [5] Katrina Nesterenko and **Lewis, Rory**. Rough sets for intelligence on embedded systems. In *International Symposium on Methodologies for Intelligent Systems*, pages 230–239. Springer, 2022.
- [6] NVIDIA Corporation. *NVIDIA CUDA C++ Programming Guide*. NVIDIA Corporation, version 11.0 edition, 2020.
- [7] Katharina Janzakova, Ismail Balafrej, Abhishek Kumar, Nishant Garg, Camille Scholaert, Jean Rouat, Dominique Drouin, Yannick Coffinier, Sébastien Pecqueur, and Fabien Alibart. Structural plasticity for neuromorphic networks with electropolymerized dendritic PEDOT connections. *Nature Communications*, 14(1):7986, 2023.
- [8] Lyes Khacef, Laurent Rodriguez, and Benoît Miramond. Brain-inspired self-organization with cellular neuromorphic computing for multimodal unsupervised learning. *Electronics*, 9(10):1605, 2020.
- [9] **Lewis, Rory**, Michael Bihn, Zhenqi Liu, and Daniel S. Barbotko. A self-learning neuromorphic system. In *The 9th International Conference on Neuroscience and Cognitive Brain Information (BRAININFO 2024)*, Athens, Greece, March 2024.
- [10] **Lewis, Rory** and Michael Bihn. Driving factor of brain growth is the connectome development. In *The 8th International Conference on Neuroscience and Cognitive Brain Information*, Barcelona, Spain, 2023.
- [11] Michael Bihn and **Lewis, Rory**. A postulate: Connectome development is the driving factor of brain growth. In *BRAININFO 2023: The Eighth International Conference on Neuroscience and Cognitive Brain Information*, pages 12–15. ThinkMind Digital Library, 2023. Best Paper Award.
- [12] Michael Bihn and **Lewis, Rory**. Synthesis of neonate connectomes for artificial sentience and common sense. In *Proceedings of the Sixth International Conference on Neuroscience and Cognitive Brain Information (BRAININFO 2021)*, pages 1–10, Nice, France, July 2021. IARIA / ThinkMind Digital Library.
- [13] Peter R. Huttenlocher. Synaptic density in human frontal cortex: Developmental changes and effects of aging. *Brain Research*, 163(2):195–205, 1979.
- [14] Pasko Rakic, Jean-Pierre Bourgeois, Mary F. Eckenhoff, Nada Zecevic, and Patricia S. Goldman-Rakic. Concurrent overproduction of synapses in diverse regions of the primate cerebral cortex. *Science*, 232(4747):232–235, 1986.
- [15] Jean-Pierre Bourgeois, Peter J. Jastreboff, and Pasko Rakic. Synaptogenesis in visual cortex of normal and preterm monkeys: Evidence for intrinsic regulation of synaptic overproduction. *Proceedings of the National Academy of Sciences of the United States of America*, 86(11):4297–4301, 1989.
- [16] **Lewis, Rory**, Michael Bihn, and Katrina Nesterenko. Rough sets for a neuromorphic cmos system. In *International Symposium on Methodologies for Intelligent Systems*, pages 102–111. Springer, 2024.

- [17] **Lewis, Rory.** The case for neuromorphic sentience to control large-scale artificial intelligence. In *Proceedings of the 10th World Conference on Information Communication Systems, Software, Security and Sustainability (WorldS4 2026)*, London, UK, July 2026. July 28–30, 2026.
- [18] Aaron T. Beck. *Depression: Clinical, Experimental, and Theoretical Aspects*. Harper & Row, 1967.
- [19] Albert Bandura. *Aggression: A Social Learning Analysis*. Prentice-Hall, 1973.
- [20] Michael E. McCullough, Robert Kurzban, and Benjamin A. Tabak. Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences*, 36(1):1–15, 2013.
- [21] Edwin H. Sutherland. *Principles of Criminology*. J. B. Lippincott, 1947.
- [22] Morris Rosenberg. *Society and the Adolescent Self-Image*. Princeton University Press, 1965.
- [23] David M. Clark. A cognitive approach to panic. *Behaviour Research and Therapy*, 24(4):461–470, 1986.
- [24] Daniel Freeman and Jason Freeman. *Paranoia: The 21st Century Fear*. Oxford University Press, 2008.
- [25] Stanley Rachman. Part i. unwanted intrusive cognitions. *Advances in Behaviour Research and Therapy*, 3(3):89–99, 1981.
- [26] Paul M. Salkovskis. Obsessional-compulsive problems: A cognitive-behavioural analysis. *Behaviour Research and Therapy*, 23(5):571–583, 1985.
- [27] Karl Jaspers. *General Psychopathology*. Johns Hopkins University Press, 1913.
- [28] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zach DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 32, 2019.
- [29] Wilson Snyder. Verilator and systemperl. In *North American SystemC Users' Group, Design Automation Conference*, 2004.
- [30] David Shah, Eddie Hung, Clifford Wolf, Serge Bazanski, Dan Gisselquist, and Miodrag Milanović. Yosys+nextpnr: an open source framework from verilog to bitstream for commercial fpgas. *arXiv preprint arXiv:1903.10407*, 2019.
- [31] Sagar Karandikar, David Biancolin, Kangning Huang, Donggyu Basu, Andrew Lim, Yakun Sophia Lee, and Krste Asanović. Firesim: Fpga-accelerated cycle-exact scale-out system simulation in the public cloud. In *Proceedings of the 45th Annual International Symposium on Computer Architecture (ISCA)*, pages 29–42, 2018.
- [32] Synopsys Inc. *Synopsys ZeBu Server Emulation System Reference Manual*. Synopsys Inc., 2024.
- [33] Cadence Design Systems. *Cadence Palladium Z1/Z2 Enterprise Emulation Platform*. Cadence Design Systems, 2023.
- [34] Aldec, Inc. *HES-DVM Hybrid Verification Platform*. Aldec, Inc., 2026. FPGA-based hardware emulation and prototyping platform.

- [35] Synopsys, Inc. *HAPS Pre-Silicon Prototyping System*. Synopsys, Inc., 2026. FPGA-based pre-silicon prototyping platform.
- [36] Cadence Design Systems, Inc. *Protium Enterprise Prototyping Platform*. Cadence Design Systems, Inc., 2026. Enterprise FPGA-based prototyping platform.
- [37] John Stickley, Deepak Kumar Garg, Brian Bailey, Jaekwang Lee, Amy Lim, Per Bojsen, Ramesh Chandra, and Ajeya Prabhakar. Understanding the accellera sce-mi transaction pipes. *IEEE Design & Test of Computers*, 29(1):32–43, 2012.
- [38] IEEE Standard for Standard SystemC Language Reference Manual, September 2023.
- [39] Didier Dubois and Henri Prade. Rough fuzzy sets and fuzzy rough sets. *International Journal of General Systems*, 17(2-3):191–209, 1990.
- [40] Yiyu Y. Yao. A comparative study of fuzzy sets and rough sets. *Information Sciences*, 109(1–4):227–242, 1998.
- [41] Katrina Nesterenko and Rory Lewis. Rough sets for intelligence on embedded systems. In *International Symposium on Methodologies for Intelligent Systems*, pages 230–239. Springer, 2022.
- [42] **Lewis, Rory**, Chad A. Mello, Yanyan Zhuang, Martin K.-C. Yeh, Yu Yan, and Dan Gopstein. Rough sets: Visually discerning neurological functionality during thought processes. pages 32–41, 2018.
- [43] **Lewis, Rory**. A semantic calculus: Common sense reasoning for information systems. pages 19th–21st September, British Computer Society at London Southbank, London, UK, 2019.
- [44] **Lewis, Rory**. A logical theory towards an invariant phenomenological state of artificial consciousness. pages Best Presentation Award, 22nd–23rd April, New York, NY, 2019.
- [45] Michael E. Bratman. *Intention, Plans, and Practical Reason*. Harvard University Press, Cambridge, MA, 1987.
- [46] Anand S. Rao and Michael P. Georgeff. Bdi agents: From theory to practice. *Technical Note 48, Australian Artificial Intelligence Institute*, pages 312–319, 1995.
- [47] Michael Wooldridge. *Reasoning about Rational Agents*. The MIT Press, Cambridge, MA, 2000.
- [48] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning About Knowledge*. The MIT Press, Cambridge, MA, 2003.
- [49] Stacy C. Marsella and Jonathan Gratch. Ema: A process model of emotion appraisal and coping. *Cognitive Systems Research*, 10(1):70–90, 2009.
- [50] Peter Spirtes, Clark N. Glymour, and Richard Scheines. *Causation, Prediction, and Search*. The MIT Press, Cambridge, MA, 2nd edition, 2000.