Proof-of-Attention (PoA):
A Human-Centric Consensus and Tokenization Mechanism

Part I: Core Mechanism and Architecture
With Illustrative TikZ/PGFPlots Figures

OCTA Research

Abstract

We introduce a blockchain consensus and tokenization mechanism, *Proof-of-Attention* (PoA), in which block rewards are minted as a function of aggregate human attention rather than computational work or stake. Time is discretized into fixed-length epochs ("attention rounds"), within which a population of devices emits bounded, entropy-rich interaction signals. These signals are filtered through a quality and anti-Sybil pipeline and aggregated into a global attention measure that determines block weight, block-level non-fungible assets, and reward allocation. This Part I develops a formal mathematical model of the PoA mechanism, including the attention field, token supply dynamics, reward functions, basic security properties, an information-theoretic view of attention, and a layered architecture. Multiple TikZ/PGFPlots figures provide visual intuition for the core constructions.

Contents

1 Global Time and Participants

1.1 Time Model

Definition 1.1 (Time slots)**.** *Let $\Delta t > 0$ denote a fixed slot duration (e.g., $\Delta t = 10$ seconds). We index slots by $t \in \mathbb{N} = \{0, 1, 2, \dots\}$. Slot t corresponds to the time interval*

$$I_t := [t\Delta t, (t+1)\Delta t).$$

Each slot t is associated with a unique candidate block B_t in the PoA chain.

1.2 Devices and Sessions

Definition 1.2 (Devices and sessions)**.** *Let $\mathcal{D}$ denote the set of devices participating in the PoA network. A* session *is an ordered pair (d, t) where $d \in \mathcal{D}$ and $t \in \mathbb{N}$ such that device d is online and eligible to emit attention signals during slot t. We denote by*

$$\mathcal{D}_t \subseteq \mathcal{D}$$

the subset of devices active in slot t.

For simplicity, we treat devices as pseudonymous entities; we explicitly avoid mapping devices to real-world identities.

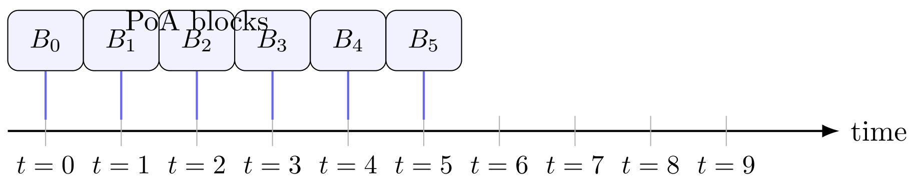

Figure 1: Discretized time into slots I_t with associated PoA blocks B_t.

1.3 Illustration: Time Slots and Blocks

2 Attention Signals

2.1 Raw Attention Trajectories

Within each slot t, each active device $d \in \mathcal{D}_t$ produces a time-resolved interaction trajectory.

Definition 2.1 (Raw attention signal)**.** *For each (d, t) with $d \in \mathcal{D}_t$, define the raw attention signal*

$$x_{d,t} : I_t \to \mathbb{R}$$

representing instantaneous interaction intensity (e.g., tap pressure proxy, touch movement magnitude, reaction strength). We assume

$$x_{d,t}(s) \geq 0, \qquad \forall s \in I_t.$$

Definition 2.2 (Raw attention mass)**.** *The raw attention mass of device d in slot t is*

$$\mathsf{RawAtt}_{d,t} := \int_{I_t} x_{d,t}(s)\,\mathrm{d}s. \tag{1}$$

2.2 Device-Level Constraints

To prevent unbounded contribution from a single device per slot, we impose per-slot and sliding-window constraints.

Definition 2.3 (Per-slot cap)**.** *Fix $A_{\max} > 0$. For every $d \in \mathcal{D}_t$, define the capped raw attention*

$$\mathsf{RawAtt}^{(\mathrm{cap})}_{d,t} := \min\{\mathsf{RawAtt}_{d,t}, A_{\max}\}. \tag{2}$$

Definition 2.4 (Sliding-window fatigue constraint)**.** *Let $W \in \mathbb{N}$ be a window length (in slots) and $C_W > 0$ be a window cap. For device d at slot t, define the cumulative capped attention in a trailing window:*

$$S^{(W)}_{d,t} := \sum_{\tau=\max\{0,t-W+1\}}^{t} \mathsf{RawAtt}^{(\mathrm{cap})}_{d,\tau}. \tag{3}$$

We enforce

$$S^{(W)}_{d,t} \leq C_W \quad \textit{for all } d, t. \tag{4}$$

Excess above C_W is suppressed by an attenuation factor (see below).

2.3 Quality and Human-ness Scoring

Attention is meaningful only when it *originates from human users*. We formalize a device-level quality score that attempts to distinguish human motor and temporal patterns from synthetic ones.

Definition 2.5 (Local feature extractor)**.** *For each* (d, t)*, let*

$$\phi_{d,t} = \phi\big(x_{d,t}\big) \in \mathbb{R}^k$$

denote a feature vector computed from the raw signal $x_{d,t}$*. Components of* $\phi_{d,t}$ *may include:*

- *inter-tap interval statistics,*
- *micro-jitter spectral properties,*
- *touch movement entropy,*
- *reaction time histograms,*
- *device sensor cross-checks.*

Definition 2.6 (Quality (human-ness) score)**.** *Given a classifier or scoring function*

$$f_{\text{qual}} : \mathbb{R}^k \to [0, 1],$$

we define the quality score for device d *in slot* t *as*

$$\mathsf{Q}_{d,t} := f_{\text{qual}}(\phi_{d,t}). \tag{5}$$

Here $\mathsf{Q}_{d,t} \approx 1$ *indicates high confidence that the interaction is consistent with human behavior, while* $\mathsf{Q}_{d,t} \approx 0$ *indicates likely synthetic or low-quality input.*

Definition 2.7 (Fatigue attenuation)**.** *Define an attenuation function* $\alpha : [0, \infty) \to (0, 1]$ *that is non-increasing in its argument and obeys:*

$$\alpha(0) = 1, \qquad \lim_{s \to \infty} \alpha(s) = 0. \tag{6}$$

For example, one may choose

$$\alpha(s) = \frac{1}{1 + (s/C_W)^\gamma} \tag{7}$$

for some $\gamma > 0$*.*

2.4 Effective Attention per Device

Definition 2.8 (Effective per-device attention)**.** *The effective attention contribution of device* d *in slot* t *is defined as*

$$\mathsf{Att}_{d,t} := \mathsf{RawAtt}^{(\text{cap})}_{d,t} \cdot \mathsf{Q}_{d,t} \cdot \alpha\big(S^{(W)}_{d,t}\big). \tag{8}$$

Remark 2.9. *Equation* (8) *encodes:*

- *a hard cap on per-slot raw attention,*
- *a soft quality scaling via* $\mathsf{Q}_{d,t}$*,*
- *a fatigue-like term* $\alpha(S^{(W)}_{d,t})$ *penalizing excessive continuous activity.*

This structure discourages botting and attention spamming across slots.

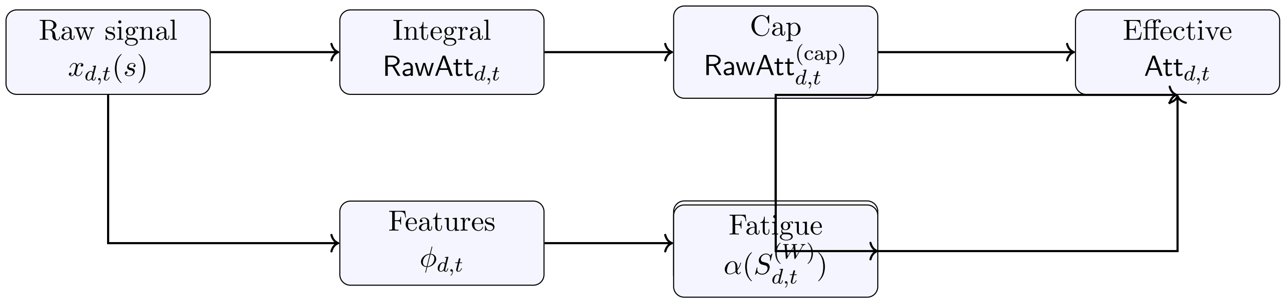

Figure 2: Device-level pipeline from raw interaction to effective attention contribution $\mathsf{Att}_{d,t}$.

2.5 Illustration: Device-Level Pipeline

3 Global Attention Field and Block Weight

3.1 Aggregate Attention per Slot

Definition 3.1 (Total effective attention per slot)**.** *The total effective attention in slot t is*

$$\mathsf{Att}_t := \sum_{d \in \mathcal{D}_t} \mathsf{Att}_{d,t}. \tag{9}$$

Definition 3.2 (Normalized attention intensity)**.** *Let $\kappa > 0$ be a scaling parameter. The normalized attention intensity for slot t is*

$$\mathsf{NormAtt}_t := \frac{\mathsf{Att}_t}{\kappa}. \tag{10}$$

3.2 Block Weight Function

Definition 3.3 (Block weight)**.** *Let $g : [0, \infty) \to [0, \infty)$ be a continuous, increasing function (e.g., concave to avoid runaway growth). We define the weight of block B_t as*

$$w_t := g(\mathsf{NormAtt}_t). \tag{11}$$

Remark 3.4. *The weight w_t may be used to modulate:*

- *token rewards allocated at slot t,*
- *rarity level of the NFT associated with block B_t,*
- *visibility and ranking of B_t in user interfaces.*

3.3 Illustration: Block Weight Functions

4 Token Supply Dynamics

4.1 Base Mint Schedule

Definition 4.1 (State variables)**.** *Let S_t denote the circulating token supply immediately* before *processing slot t. Let λ_t denote any exogenous parameters at time t (e.g., epoch index, halving schedule phase).*

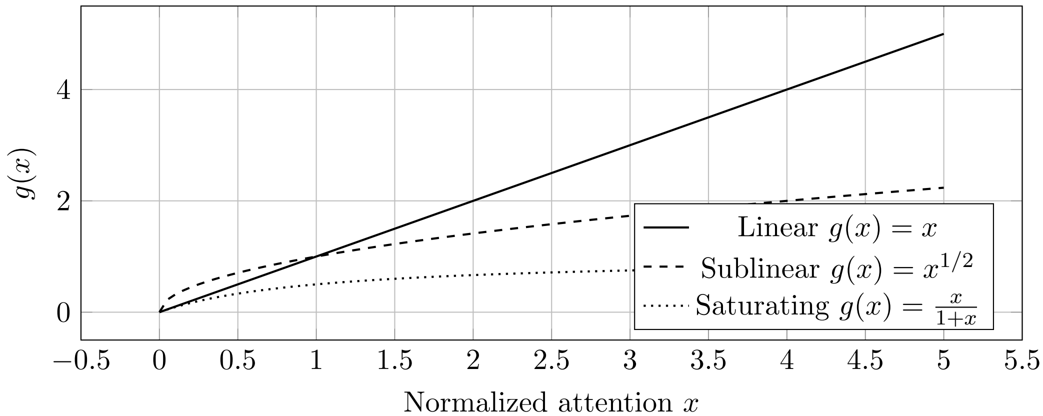

Figure 3: Illustrative choices for the block weight function $g(x)$ as attention increases.

Definition 4.2 (Base mint function)**.** *Define a base mint function*

$$M_{\text{base}} : [0, \infty) \times [0, \infty) \times \Lambda \to [0, \infty),$$

where Λ is the space of exogenous parameters, such that

$$M_{\text{base}}(w_t, S_t; \lambda_t) \tag{12}$$

is the baseline newly minted amount of tokens for slot t before burn and bonus adjustments.

A simple example is

$$M_{\text{base}}(w_t, S_t; \lambda_t) = \mu(\lambda_t) \frac{w_t}{1 + S_t/K}, \tag{13}$$

where $\mu(\cdot)$ is a schedule-dependent scale, and $K > 0$ is an "effective capacity" controlling diminishing emissions as S_t grows.

4.2 Attention-Responsive Burn

To self-stabilize in periods of extremely high attention, we can include a burn component that increases faster than the mint when Att_t exceeds an equilibrium range.

Definition 4.3 (Burn function)**.** *Let $h : [0, \infty) \to [0, \infty)$ be a convex, non-decreasing function with $h(0) = 0$. We define the burn amount at slot t as*

$$B_t := h\big(\mathsf{NormAtt}_t\big). \tag{14}$$

A canonical example:

$$h(x) = \eta\, [x - x_0]_+^2, \qquad [z]_+ = \max\{z, 0\}, \tag{15}$$

where x_0 is a target normalized attention and $\eta > 0$ is a burn sensitivity parameter.

4.3 Net Mint and Supply Recurrence

Definition 4.4 (Net mint per slot)**.** *The net minted quantity at slot t is*

$$M_t := M_{\text{base}}(w_t, S_t; \lambda_t) - B_t. \tag{16}$$

We constrain the mechanism such that $M_t \geq 0$ for all t (e.g., by taking $M_t := \max\{0, M_t\}$).

Definition 4.5 (Supply recurrence)**.** *The token supply dynamics are given by*

$$S_{t+1} = S_t + M_t. \tag{17}$$

Remark 4.6. *Equation (17) defines a discrete-time dynamical system on $\mathbb{R}_{\geq 0}$ driven by the attention sequence (Att_t) and the policy functions g, M_{base}, h.*

4.4 Reward Allocation to Participants

Definition 4.7 (Per-device reward share)**.** *For $d \in \mathcal{D}_t$ with $\mathsf{Att}_{d,t} > 0$, define the base share*

$$r^{(0)}_{d,t} := \frac{\mathsf{Att}_{d,t}}{\mathsf{Att}_t}. \tag{18}$$

If $\mathsf{Att}_t = 0$, no rewards are minted, i.e., $M_t = 0$.

Definition 4.8 (Role multipliers)**.** *Let $m_{d,t} \geq 0$ denote a multiplicative role factor for device d in slot t encoding roles such as:*

- initiator *(first signalling device),*
- amplifier *(drives secondary waves of attention),*
- closer *(sustains attention near the end of the slot),*
- streaker *(maintains an attention streak across slots).*

We define the normalized role multiplier

$$\tilde{m}_{d,t} := \frac{m_{d,t}}{\sum_{d' \in \mathcal{D}_t} m_{d',t} r^{(0)}_{d',t}}, \tag{19}$$

with the convention that if the denominator is 0, then $\tilde{m}_{d,t} := 1$ for all d.

Definition 4.9 (Final per-device reward)**.** *The final token reward for device d at slot t is*

$$\mathsf{Reward}_{d,t} := M_t \, r^{(0)}_{d,t} \, \tilde{m}_{d,t}. \tag{20}$$

Proposition 4.10 (Reward conservation)**.** *For each slot t, the reward allocation is supply-preserving in the sense that*

$$\sum_{d \in \mathcal{D}_t} \mathsf{Reward}_{d,t} = M_t.$$

Proof. By definition,

$$\begin{aligned}
\sum_{d \in \mathcal{D}_t} \mathsf{Reward}_{d,t} &= M_t \sum_{d \in \mathcal{D}_t} r^{(0)}_{d,t} \tilde{m}_{d,t} \\
&= M_t \sum_{d \in \mathcal{D}_t} r^{(0)}_{d,t} \frac{m_{d,t}}{\sum_{d'} m_{d',t} r^{(0)}_{d',t}} \\
&= M_t \frac{\sum_d m_{d,t} r^{(0)}_{d,t}}{\sum_{d'} m_{d',t} r^{(0)}_{d',t}} \\
&= M_t.
\end{aligned}$$

□

4.5 Illustration: Burn Function

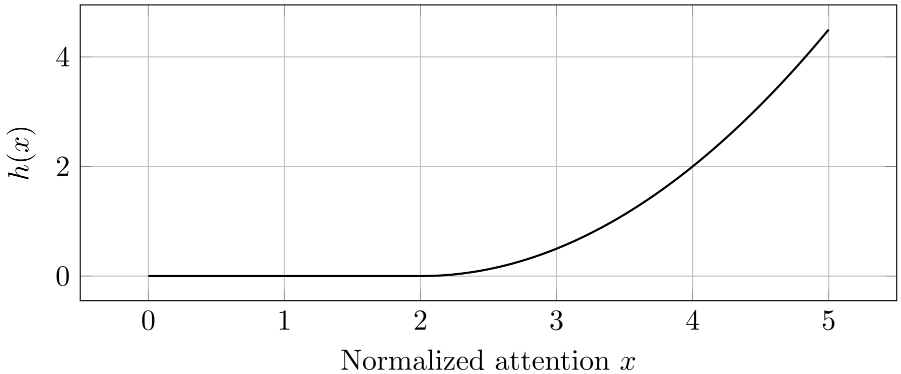

Figure 4: Example burn function $h(x) = \eta[x - x_0]_+^2$ with $x_0 = 2$, $\eta = 0.5$.

5 NFT Metadata for Blocks

Each block B_t can be associated with a non-fungible token (NFT) that encodes attention-based metadata.

Definition 5.1 (Attention snapshot vector)**.** *Let $\boldsymbol{\xi}_t$ denote a finite-dimensional vector of aggregate block-level statistics computed from $\{x_{d,t}\}_{d\in\mathcal{D}_t}$ and/or $\{\mathsf{Att}_{d,t}\}_{d\in\mathcal{D}_t}$, including for example:*

- *total attention Att_t,*
- *number of participating devices $|\mathcal{D}_t|$,*
- *entropy of device-level contributions,*
- *temporal concentration of attention within I_t,*
- *sentiment or reaction composition.*

Definition 5.2 (Block NFT)**.** *To each block B_t we associate an NFT with metadata*

$$\mathrm{Meta}(B_t) := (t, w_t, \boldsymbol{\xi}_t, H_t),$$

where H_t is the block header hash including canonical references to PoA state.

Definition 5.3 (Rarity score)**.** *Define a rarity function $R : \mathbb{R}^m \to \mathbb{R}_{\geq 0}$ such that*

$$\rho_t := R(\boldsymbol{\xi}_t)$$

is the rarity score of block B_t. For example, R can be designed as a combination of:

- *extremal attention events (e.g., high Att_t),*
- *rare temporal patterns (e.g., bursts at specific times),*
- *concentration metrics (e.g., near-unanimous device participation).*

6 Basic Security and Anti-Sybil Considerations

We sketch a simplified model for Sybil resistance.

6.1 Adversarial Model

Definition 6.1 (Adversary model)**.** *An adversary controls a subset $\mathcal{D}^{\mathrm{adv}} \subset \mathcal{D}$ of devices (or device-like entities). The adversary aims to:*

- *maximize its share of Att_t in each slot,*
- *pass the quality filter to obtain high $\mathsf{Q}_{d,t}$,*
- *circumvent fatigue constraints by splitting effort across Sybil devices.*

6.2 Effective Attention Under Attack

Definition 6.2 (Adversarial and honest attention)**.** *Let*

$$\mathsf{Att}_t^{\mathrm{H}} := \sum_{d \in \mathcal{D}_t \setminus \mathcal{D}^{\mathrm{adv}}} \mathsf{Att}_{d,t}, \quad \mathsf{Att}_t^{\mathrm{A}} := \sum_{d \in \mathcal{D}_t \cap \mathcal{D}^{\mathrm{adv}}} \mathsf{Att}_{d,t}. \tag{21}$$

Thus, $\mathsf{Att}_t = \mathsf{Att}_t^{\mathrm{H}} + \mathsf{Att}_t^{\mathrm{A}}$.

Definition 6.3 (Adversarial share)**.** *In slot t, the adversarial share of effective attention is*

$$\sigma_t := \frac{\mathsf{Att}_t^{\mathrm{A}}}{\mathsf{Att}_t} \in [0, 1], \tag{22}$$

whenever $\mathsf{Att}_t > 0$.

The goal is to ensure that, under reasonable assumptions on the quality function $\mathsf{Q}_{d,t}$ and attenuation α, it is costly or difficult for σ_t to approach 1 persistently.

6.3 Qualitative Bound

Proposition 6.4 (Informal adversarial bound)**.** *Suppose:*

(i) For honest devices, $\mathbb{E}[\mathsf{Q}_{d,t}] \approx q_{\mathrm{H}}$ with q_{H} bounded away from 0.

(ii) For adversarial devices attempting to emulate human patterns, any strategy that increases $\mathsf{RawAtt}_{d,t}^{(\mathrm{cap})}$ beyond typical human ranges leads to either reduced $\mathsf{Q}_{d,t}$ or large $S_{d,t}^{(W)}$, thus smaller $\alpha(S_{d,t}^{(W)})$.

Then there exists a constant $\sigma^\star < 1$ such that, in expectation over time and randomness of the scoring process,

$$\mathbb{E}[\sigma_t] \leq \sigma^\star,$$

for any adversarial strategy that maintains bounded computational resources and does not compromise the central quality-scoring pipeline.

Remark 6.5. *A detailed proof would depend on the concrete specification of f_{qual} and α, as well as an explicit model of adversarial capabilities. The key point is that PoA leverages the difficulty of perfectly emulating the microstructure of human motor and interaction patterns at scale.*

7 Consensus and Chain Finality (Abstract View)

The PoA mechanism described thus far specifies *how* block rewards and NFTs are generated from attention. A separate consensus protocol is required to order blocks and provide finality.

Definition 7.1 (Abstract consensus layer)**.** *Let $\mathcal{C}$ denote an underlying consensus protocol (e.g., classical BFT, PoS-based finality, or a committee-based protocol) which:*

- *orders candidate blocks B_t into a chain,*
- *ensures block headers commit to PoA state* $(\mathsf{Att}_t, w_t, \boldsymbol{\xi}_t)$*,*
- *enforces the supply recurrence* (17)*.*

We assume that $\mathcal{C}$ guarantees standard safety and liveness properties under its own assumptions.

Remark 7.2. *PoA can be viewed as an* attention-weighted reward and metadata layer *coupled to a conventional consensus engine. Alternatively, the consensus engine itself can incorporate PoA-derived weights in block selection, where block proposal probability is proportional to attention-weighted stake. Both constructions are compatible with the mathematical framework above.*

7.1 Illustration: Layered Architecture

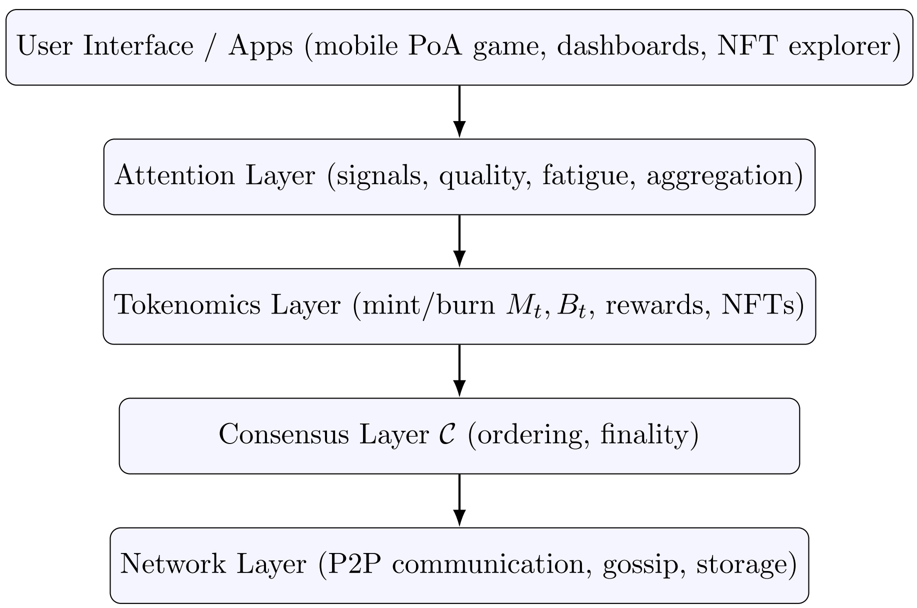

Figure 5: Conceptual architecture: PoA attention and tokenomics layers sit above a generic consensus protocol and network layer.

8 Information-Theoretic View of Attention

We can refine PoA further by treating attention as an information-bearing stochastic process and tying rewards to *surprise* and *diversity* in the attention field.

8.1 Attention as a Distribution

Definition 8.1 (Attention distribution over devices)**.** *For a given slot t with $\mathsf{Att}_t > 0$, define the device-level attention distribution*

$$p_{d,t} := \frac{\mathsf{Att}_{d,t}}{\mathsf{Att}_t}, \qquad d \in \mathcal{D}_t.$$

Definition 8.2 (Attention entropy)**.** *The Shannon entropy of the attention distribution at slot t is*

$$H_t := -\sum_{d \in \mathcal{D}_t} p_{d,t} \log p_{d,t}.$$

Remark 8.3. *High entropy corresponds to a more decentralized attention pattern (many devices engaged), while low entropy corresponds to concentrated attention (few devices dominate).*

8.2 Entropy-Aware Reward Adjustment

Definition 8.4 (Entropy-based modifier)**.** *Let $\psi : [0, \infty) \to (0, \infty)$ be a continuous function. Define an entropy-dependent multiplier*

$$\chi_t := \psi(H_t).$$

We may redefine the base mint as

$$M_{\text{base}}^{(\text{ent})}(w_t, S_t; \lambda_t) := \chi_t \, M_{\text{base}}(w_t, S_t; \lambda_t),$$

thus rewarding not only total *attention but also its* dispersion*.*

8.3 Illustration: Entropy Modulation

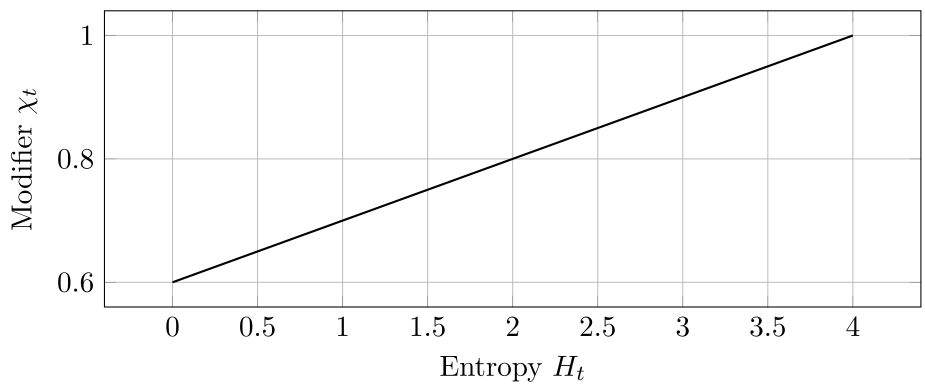

Figure 6: Example of an entropy-based modifier χ_t that increases linearly with attention entropy, encouraging broader, more decentralized participation.

9 Implementation Notes: On-Chain vs Off-Chain PoA

Here we outline a practical separation of concerns between off-chain attention aggregation and on-chain verification.

9.1 Off-Chain Aggregation

- Mobile clients collect raw interaction traces $x_{d,t}(s)$.
- Local feature extraction and partial scoring produce $\phi_{d,t}$ and preliminary $\mathsf{Q}_{d,t}$.
- Aggregator nodes collect commitments to $(\mathsf{RawAtt}_{d,t}^{(\text{cap})}, \mathsf{Q}_{d,t})$, possibly via secure enclaves or zero-knowledge proofs.
- Aggregators compute Att_t, H_t, and associated statistics $\boldsymbol{\xi}_t$.

9.2 On-Chain Verification and Settlement

- Aggregators submit a succinct proof (e.g., zk-SNARK) that:
  - all device contributions obey caps and fatigue bounds,
  - quality scores are within allowed ranges,
  - the aggregate values $\mathsf{Att}_t, H_t, \boldsymbol{\xi}_t$ are computed correctly.
- Smart contracts:
  - verify proofs,
  - compute w_t, M_t, B_t,
  - update supply S_{t+1},
  - distribute $\mathsf{Reward}_{d,t}$ to participants (or to aggregator contracts representing them),
  - mint NFTs associated with B_t.

9.3 Illustration: Off-Chain / On-Chain Split

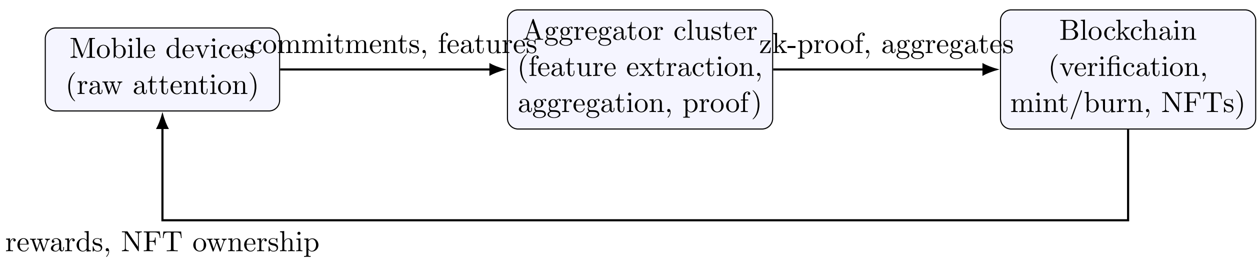

Figure 7: High-level implementation flow: off-chain attention processing with on-chain cryptographic verification and settlement.

10 Summary of Part I

We have formalized a human-centric tokenization mechanism, Proof-of-Attention (PoA), as a discrete-time dynamical system driven by:

- device-level attention signals with caps and fatigue,
- quality scores distinguishing human from synthetic interactions,
- an attention aggregation and block-weight function,
- supply and burn functions responsive to attention intensity,
- attention-proportional reward allocation with role multipliers,
- entropy-aware dispersion rewards and a practical off-chain/on-chain split.

Part II develops the full game-theoretic, Bayesian, and evolutionary analysis of PoA as a repeated stochastic game, and studies incentive compatibility, coalitions, and welfare alignment under different policy choices.