Proof-of-Attention (PoA):
Game-Theoretic and Information-Theoretic Analysis

Part II-A: Strategic Foundations, Best Responses, and Coalitions

OCTA Research

Abstract

Part II-A develops the foundational game-theoretic model of Proof-of-Attention (PoA). We define players, actions, attention costs, utilities, and repeated-game payoffs. We analyze best responses, Nash equilibria, Sybil splitting, human–bot substitution, and coalition stability. TikZ and PGFPlots figures visualize payoff landscapes and coalition benefits in computationally lightweight forms.

Contents

1 Players, Actions, Utilities

Definition 1.1 (Players)**.** *Let $\mathcal{P}$ be the set of players, each controlling one or more devices.*

Each player i selects $a_{i,t}$ producing effective attention and cost $c_i(a_{i,t})$.

Definition 1.2 (Reward and Utility)**.** *Let M_t denote the net mint in slot t, and $\mathsf{Att}_{i,t}$ the effective attention of player i with total $\mathsf{Att}_t = \sum_{j\in\mathcal{P}} \mathsf{Att}_{j,t}$. Then*

$$R_{i,t} = M_t \frac{\mathsf{Att}_{i,t}}{\mathsf{Att}_t}, \qquad U_{i,t} = v(R_{i,t}) - c_i(a_{i,t}),$$

where v is a (possibly concave) valuation function.

The lifetime utility of player i for discount factor $0 < \delta < 1$ is

$$U_i = \sum_{t=0}^{\infty} \delta^t U_{i,t}.$$

2 Static PoA Game and Nash Equilibrium

Fix a slot t and treat M_t as exogenous.

Definition 2.1 (Static PoA game for slot t). *Players choose $a_{i,t}$ inducing $\mathsf{Att}_{i,t}$ with*

$$\sum_{i \in \mathcal{P}} \mathsf{Att}_{i,t} = \mathsf{Att}_t.$$

Player i's share is $s_{i,t} = \mathsf{Att}_{i,t}/\mathsf{Att}_t$ and reward $R_{i,t} = M_t s_{i,t}$, utility $U_{i,t} = v(M_t s_{i,t}) - c_i(a_{i,t})$.

Definition 2.2 (Best response). *Given others' actions $a_{-i,t}$, a best response of player i is*

$$a^\star_{i,t} \in \arg\max_{a_{i,t}} [U_{i,t} \mid a_{i,t}, a_{-i,t}].$$

Definition 2.3 (Nash equilibrium). *An action profile $(a^\star_{i,t})_i$ is a Nash equilibrium if each $a^\star_{i,t}$ is a best response to $a^\star_{-i,t}$.*

Proposition 2.4 (Symmetric equilibrium). *If v is strictly concave, and c_i is strictly convex with identical functional form for all i, then there exists a symmetric Nash equilibrium where all participating players choose identical effective attention $\mathsf{Att}^\star_{i,t}$.*

Remark 2.5. *Intuitively, at the symmetric equilibrium, marginal reward equals marginal cost for each player; higher attention than equilibrium reduces net utility due to steeply increasing cost, while lower attention sacrifices reward share.*

3 Illustrative Payoff Landscape (Lightweight)

The original 3D surface can be approximated by a 2D cross-section that is much lighter to render. Fix Player 2's effort a_2 and consider Player 1's utility as a function of a_1:

$$U_1(a_1 \mid a_2) = \frac{a_1}{a_1 + a_2} - \gamma a_1^2,$$

for some $\gamma > 0$.

4 Sybil Splitting

4.1 Splitting Across Pseudonymous Identities

Definition 4.1 (Sybil split). *Suppose player i uses k pseudonymous identities, splitting total effective attention $\mathsf{Att}_{i,t}$ into $\mathsf{Att}^{(1)}_{i,t}, \ldots, \mathsf{Att}^{(k)}_{i,t}$ with*

$$\sum_{j=1}^{k} \mathsf{Att}^{(j)}_{i,t} = \mathsf{Att}_{i,t}.$$

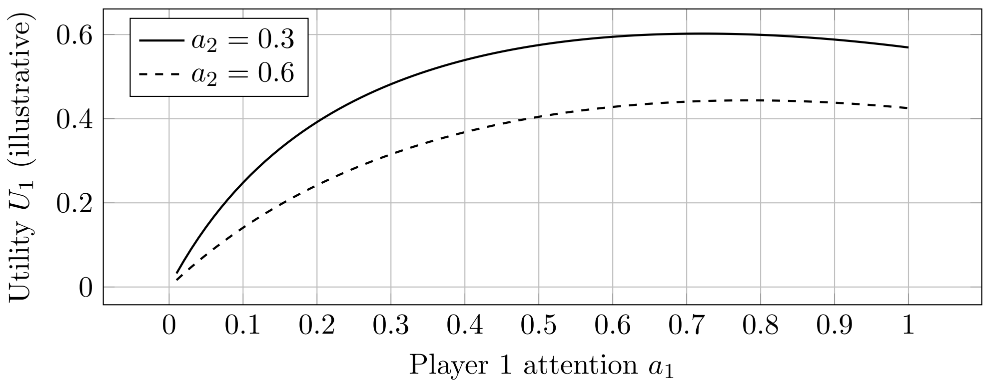

Figure 1: Player 1 payoff cross-sections for two fixed values of Player 2's attention. This 2D plot replaces the heavier 3D surface but preserves the qualitative shape (interior optimum where marginal benefit = marginal cost).

If quality and fatigue scores were independent of splitting, total reward would be

$$R_{i,t} = \sum_{j=1}^{k} M_t \frac{\mathsf{Att}_{i,t}^{(j)}}{\mathsf{Att}_t} = M_t \frac{\mathsf{Att}_{i,t}}{\mathsf{Att}_t}.$$

Proposition 4.2 (Sybil irrelevance under perfect proportionality)**.** *In a purely proportional model where quality and fatigue scores do not change under splitting, Sybil splitting is reward-neutral.*

Remark 4.3. *In PoA, per-device caps, fatigue constraints, and quality scoring make splitting* worse*, not neutral: attempts to simulate many identities typically reduce* Q *and increase the fatigue term, so* $\mathsf{Att}_{i,t}$ *falls for fixed raw effort.*

5 Human vs Bot Substitution

5.1 Mixed Effort Model

Let player i allocate effort between human and bot components:

$$e_{i,t}^{\mathrm{H}} \quad \text{(human effort)}, \qquad e_{i,t}^{\mathrm{B}} \quad \text{(bot effort)}.$$

Assume effective attention:

$$\mathsf{Att}_{i,t} = \mathsf{Q}^{\mathrm{H}} e_{i,t}^{\mathrm{H}} + \mathsf{Q}^{\mathrm{B}} e_{i,t}^{\mathrm{B}},$$

with $\mathsf{Q}^{\mathrm{H}} > \mathsf{Q}^{\mathrm{B}}$.

Costs:

$$c_i(a_{i,t}) = c_H(e_{i,t}^{\mathrm{H}}) + c_B(e_{i,t}^{\mathrm{B}}),$$

where c_H and c_B are increasing and convex.

Proposition 5.1 (Critical bot-substitution threshold)**.** *There exists a critical level of bot cost efficiency*

$$\theta^\star = \frac{c_H'(0)/\mathsf{Q}^{\mathrm{H}}}{c_B'(0)/\mathsf{Q}^{\mathrm{B}}}$$

such that:

- *if $\theta < \theta^\star$, then human effort dominates for marginal increases in total effort;*
- *if $\theta > \theta^\star$, then bot effort becomes attractive unless penalized by Q or fatigue α.*

Remark 5.2. *PoA design aims to push $\theta^\star$ high by:*

- *making Q^{B} significantly lower than Q^{H},*
- *tightening fatigue against bot-like patterns,*
- *raising $c_B'(0)$ via requirements for high-fidelity human emulation.*

6 Coordination, Guilds, and Coalitions

6.1 Coalition Attention and Utility

Definition 6.1 (Coalition)**.** *A coalition $C \subseteq \mathcal{P}$ is a subset of players that coordinate strategies. Total coalition attention:*

$$\mathsf{Att}_{C,t} := \sum_{i \in C} \mathsf{Att}_{i,t},$$

and coalition utility:

$$U_C := \sum_{i \in C} U_i.$$

6.2 Coalition Stability and Core

Definition 6.2 (Coalitionally rational)**.** *A coalition C is stable if there is no subgroup $C' \subseteq C$ such that all players in C' strictly increase their utilities by deviating jointly.*

The set of payoff allocations that cannot be improved upon by any coalition is the *core* of the cooperative PoA game.

Remark 6.3. *Because PoA rewards scale linearly with attention share but costs and fatigue are convex, coalitions may:*

- *share fixed costs of learning optimal engagement patterns,*
- *coordinate streak timing and role multipliers,*
- *balance members to avoid fatigue saturation.*

Coalition dominance is limited by device-level caps, quality scoring, and entropy-based dispersion incentives (from Part I).

6.3 Illustration: Coalition vs Solo Payoff

7 Conclusion of Part II-A

This part established:

- the static PoA game with attention-based reward sharing,
- best responses and symmetric Nash equilibria,

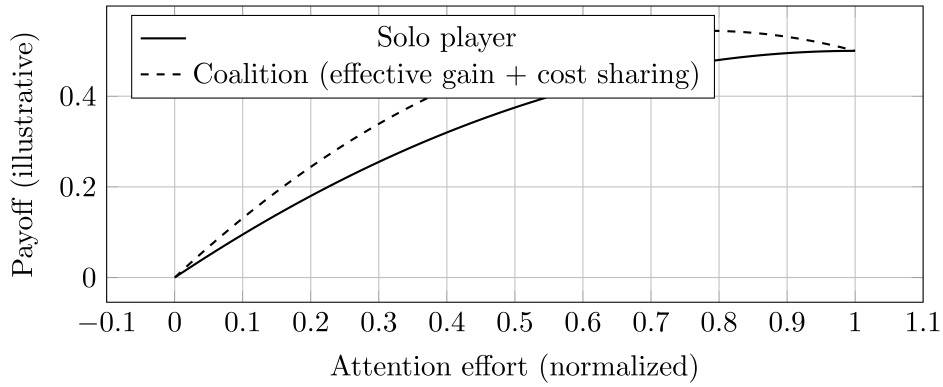

Figure 2: Illustrative payoff profiles for solo vs coalition behavior as a function of contributed effort. Coalitions can increase payoff at moderate effort but face stronger diminishing returns at high effort.

- conditions under which Sybil splitting is neutral or strictly dominated,
- a mixed human/bot effort model and a critical substitution threshold,
- coalition formation and qualitative stability via convex costs and fatigue constraints.

Part II-B extends the analysis to evolutionary dynamics, welfare alignment, Bayesian private-type incentives, and policy calibration of mint–burn curves.