BAR Nash Equilibrium and Application to Blockchain Design

Designing a Solution for the Verifier's Dilemma in Quorum-Based Blockchains

Maxime Reynouard[1,3], Olga Gorelkina[2], Rida Laraki[1,2]

1. Université Paris Dauphine - PSL
2. UM6P
3. Nomadic Labs

21 September 2023

Outline

Setup

- T the strategy space, $\tau \in T$ the prescribed protocol
- $N = \{1, 2, ..., n\}$, the set of all agents, $i \in N$ a single agent
- $s \in T^n$ a (joint) strategy profile,
- $s_i \in T$ the strategy of agent i for the profile s
- $s_I \in T^{|I|}$ the sub-profile of agents in $I \subset N$
- $u_i(s)$ the payoff of agent i
- $I, J \subset N$ disjoint, $i, j \in N \setminus (I \cup J)$, we write
  - $(s_I, s_J, s_i, s_j) = s_{I \cup J \cup \{i,j\}})$
  - $u_i(s) = u_i(s_1, ...s_n) = u_i(s_I, s_{N \setminus I})$

Symmetric games

Definition

A game is symmetric iff $u_i(s_1, \ldots, s_n) = u_{\pi(i)}(s_{\pi(1)}, \ldots, s_{\pi(n)})$ for any permutation π over N.

The BAR model: three types of agents

Following the *Byzantine–Altruistic–Rational* (*BAR*) model, we distinguish three types of agents:

- $F \subset N$, the Faulty (Byzantine) agents that deviate arbitrarily from τ. Their behaviour may range from non-strategic faults to collusive attacks.
- $G \subset N$, the Gain seeking (Rational) agents maximizing their payoff u_i.
- $H \subset N$, the Honest (Altruistic) agents following τ unconditionally
- F, G, and H partition N, so
  - they are distinct
  - their cardinals f, g, and h sum to n

BAR-Robust Equilibrium

Definition

A joint strategy profile $s^* \in T^n$ is a $(\bar{f}, \bar{g})$ **BAR-robust equilibrium** for two given integers $\bar{f}$ and $\bar{g}$ if:

1. For all $F \subset N$ such that $f \leq \bar{f}$, $s_F \in T^f$ and $i \in N \setminus F$:
   $u_i(s_F, s^*_{N \setminus F}) \geq u_i(s^*)$.
2. For all disjoint sets $F, G \subset N$, and strategy profile $s \in T^n$ such that $g \leq \bar{g}$ and $f \leq \bar{f}$, where $s_G \in T^g$ and $s_F \in T^f$, there exists $i \in G$ such that $u_i(s_F, s_G, s^*_{N \setminus (F \cup G)}) \leq u_i(s_F, s^*_{N \setminus F})$.

- (1) corresponds to byzantine fault tolerance (BFT) in the distributed computing literature
- (2) is equivalent to the *strong Nash equilibrium* condition when $g = n$.

Ittai Abraham, Lorenzo Alvisi, and Joseph Y. Halpern. "Distributed Computing Meets Game Theory: Combining Insights from Two Fields". In: *SIGACT News* 42.2 (June 2011), pp. 69–76

Drawbacks

Both conditions are fairly restrictive

- With $\bar{g} \geq 1$, condition (2) implies that s^* is a Nash equilibrium (let $f = 0$ and $g = 1$).
- With $\bar{g} \geq 2$, condition (2) further implies that no two players can jointly deviate to simultaneously increase their payoff (let $f = 0$ and $g = 2$).

Already in the prisoner's dilemma these two conditions are incompatible.

In symmetric games, conditions (1) and (2) imply that the equilibrium strategy of *Rationals* is a best reply to all possible deviations of *Byzantines*.

New Concept: BARNE

Definition

The joint strategy profile $s^*_G \in T^g$ is

1. BARNE at (F, G) with $F,\ G \subset N$ disjoint, if:
   for all $i \in G$, $s^*_i \in argmax_{s_i \in T}\ min_{s_F \in T^f}\ u_i(s_F, s_i, s^*_{G \setminus \{i\}}, s_H)$.
2. BARNE at (f, g) if:
   for all F and G such that $|F| = f$ and $|G| = g$, s^*_G is a BARNE at (F, G).

Existence of BARNE

In contrast to the BAR-robust equilibrium the BARNE is guaranteed to exist under the following conditions.

Theorem

For F and G, two disjoint subsets of N, noting $H = N \setminus (F \cup G)$, if (1) T is a convex compact subset of a topological vector space, (2) any $i \in G$, u_i is continuous and (3) $t_i \mapsto u_i(s_F, (t_i, s_{G\setminus\{i\}}), s_H)$ is concave for any strategy profile $s \in T^n$, then a BARNE exists at (F, G).
*Moreover, if the game is **symmetric** then for every (f, g) there exists a **symmetric BARNE** at (f, g) that is, $\exists \sigma \in T$ s.t. $s^*_G = \sigma^g$ is a BARNE at (f, g).*

Hence, the existence of a BARNE is guaranteed in particular in mixed extensions of finite games.

A congestion game

Example

- $T = \{A,\ B\}$
- Parameter $k \in \mathbb{N}^*,\ k < n$
- $u_i(s) = \begin{cases} 1 & \text{if } s_i = A \\ 2 & \text{if } s_i = B \text{ and } \sum_{j=1}^{n} \mathbb{1}(s_j = B) \leq k \\ 0 & \text{if } s_i = B \text{ and } \sum_{j=1}^{n} \mathbb{1}(s_j = B) > k \end{cases}$

Different sensible prescribed strategy can be imagined $\tau = A$, or even a prescribed profile with k agents playing B, the rest playing A

- In a standard game theory setting: $g = n$, $f = h = 0$: numerous equilibria, k agents play B, the rest plays A
- In the BAR model, BAR-robust equilibria are impossible:
  - Byzantines can deviate from A to B to lower payoffs (no BFT)
  - Rationals cannot best reply, A could mean missing out on $u_i = 2$ from B while B means risking 0 if their is a congestion
- Several BARNE exist, with $\tau = A$, let $max(k - f, 0)$ Rationals play B while the others safely play A

BARNE Refinement: Local Stability

Definition

A strategy $\sigma \in T$ constitutes a δ**-stable BARNE with respect to norm** $\|.\|_\nu$ **at** $(\dot{f}, \dot{g})$, if for all (f, g) such that $\left\|(\dot{f}, \dot{g}) - (f, g)\right\|_\nu \leq \delta$, σ is a symmetric BARNE at (f, g).

The choice of the relevant norm $\|.\|_\nu$ depends on the application.

- Intuitive:
  - $\|.\|_2$, Euclidean but bad when when a byzantine becomes a rational
  - $\|.\|_\infty$ non-Euclidean
- More complex, but better properties:
  $\|.\|_{2^*}$: $(f, g) \mapsto \frac{1}{\sqrt{2}}\|(f, g, n - f - g)\|_2$.

BARNE Refinement: Global Stability

Global stability, is more closely related to the notion of fault tolerance.

Definition

A strategy $\sigma \in T$ constitutes a **globally stable symmetric BARNE** at $(\bar{f}, \bar{g})$ if for all disjoint subsets F and G of N such that $f \leq \bar{f}$ and $g \leq \bar{g}$, σ is a BARNE at (F, G).

Note that in example 5, no equilibrium would be globally stable, however, when $f > k$, the equilibrium where rational agents all play A is $(f - k)$-stable. This is because even with $f - k$ less Byzantine agents, if one rational chooses B then byzantine can crash it.

Properties of Different Notions of Equilibrium

	BAR-rob.	BARNE	L.S. BARNE	G.S. BARNE
anti-coalition	✓			
anti-deviations	✓	✓	✓	✓
dominant	✓			
max-min	✓	✓	✓	✓
locally stable	✓		✓	✓
globally stable	✓			✓

A simplified Quorum-Based Consensus Protocol

```
▷ BEGINNING OF A NEW ROUND / PROPOSAL
if We are the round proposer then
    Create a new valid block b
    Propose b on the network
▷ ENDORSING
while NOT (round timeout OR endorsed this round) do
    if We receive a new block proposal B then
        Check validity of B
        if B is valid then
            Endorse B
▷ DECISION
while NOT round timeout do
    if We received Q or more endorsements for B then
        add B to our blockchain
        GO TO next level
GO TO next round
```

Rational Agent's Payoffs

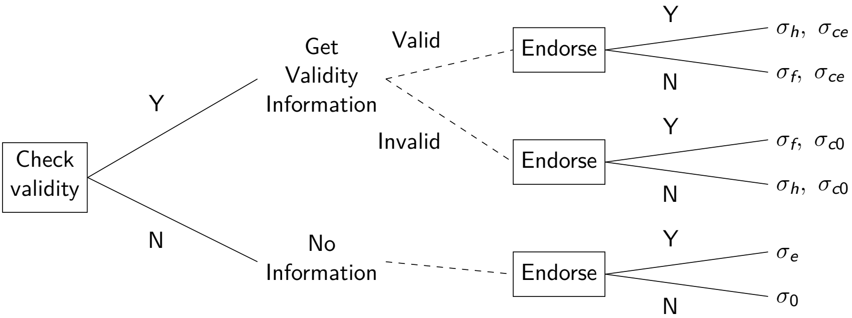
- If quorum Q is not reached
  - no Loss nor Reward
- If quorum Q is reached
  - reward r_e if endorsed
  - Loss L if Invalid
- Checking validity: c_c

Hence, a Rational agent's payoff is:

$$u = \mathbb{1}_{Accepted\ Block}(\mathbb{1}_{Endorsed}\, r_e - \mathbb{1}_{Invalid\ Block}\, L) - \mathbb{1}_{Checked\ Validity}\, c_c$$

where we assume that $L \gg r_e \gg c_c > 0$.

Decision Tree in the Endorsement Game

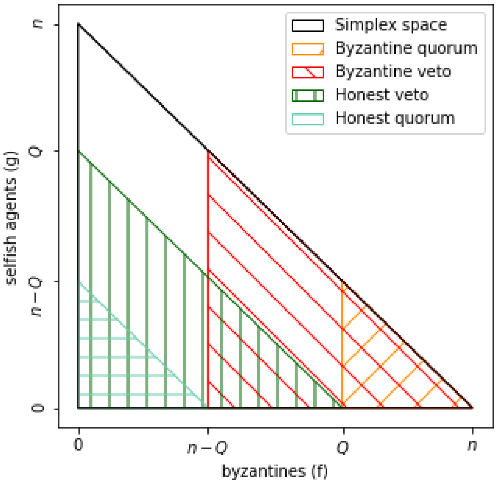


Six pure strategies

- σ_{ce}: Check validity, endorse unconditionally
- σ_{c0}: Check validity, do not endorse unconditionally
- σ_h: Check validity, endorse iff the block is valid.
  (The prescribed strategy that Honest or Altruistic agents follow.)
- σ_f: Check validity, endorse iff the block is invalid.
  (The minimising strategy of the Byzantine players.)
- σ_e: Do not check validity, endorse unconditionally
- σ_0: Do not check validity, do not endorse unconditionally

Due to dominance, *Rationals* only choose among σ_e, σ_0, and σ_h $(= \tau)$. *Byzantines* play σ_f in a symmetric BARNE.

The Byzantine-rational simplex



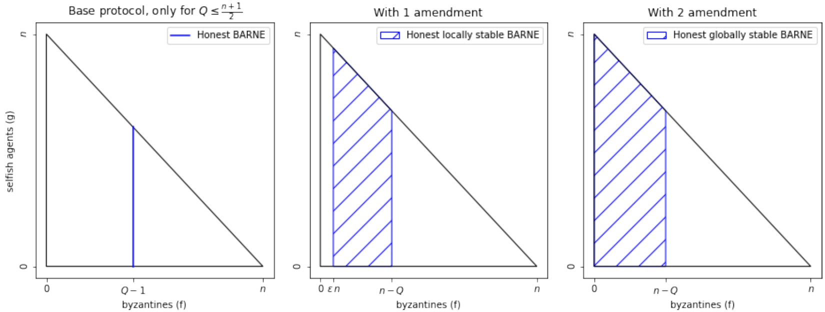
When we are both in the honest veto and honest quorum (f and g are smallish), σ_e dominates σ_h

Amending the Protocol in 2 steps

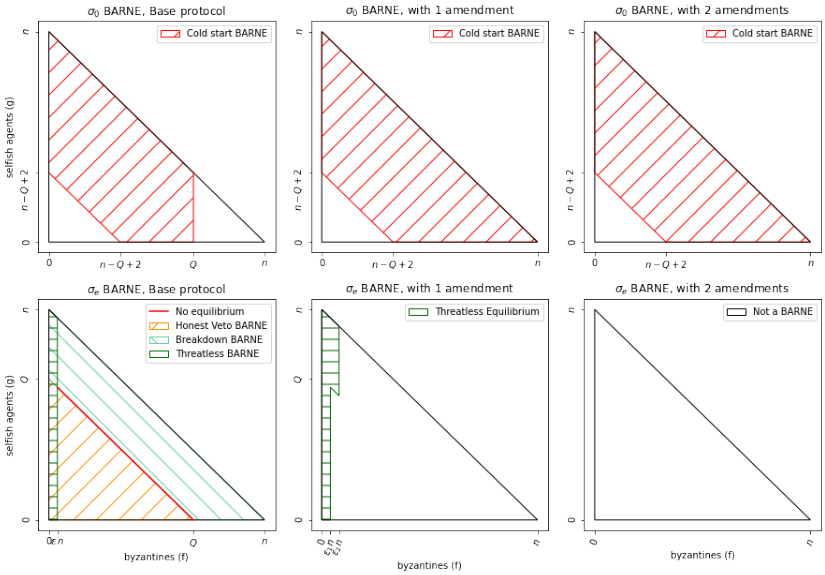
- Fine invalid block endorsers with $L_e \gg r_e$
  - Proof is transmitted, stake is slashed, possibly given to accuser
  - Insufficient when little to no Invalid block ($f \ll n$), then the fine is not a credible threat.
- Trap blocks
  - Private information gives the right to propose an invalid block
  - Ensures invalid blocks have a minimal probability of being proposed $p_{trap} > \frac{r_e + c_c}{L_e}$

Inspired by similar solution to the free-rider problem in rollups Jason Teutsch and Christian Reitwießner. "A scalable verification solution for blockchains". In: *CoRR* abs/1908.04756 (2019). arXiv: 1908.04756. URL: http://arxiv.org/abs/1908.04756

Areas where the honest strategy σ_h is a BARNE



Other strategies BARNE



Thanks

Questions ?

Payoff table

$u(\sigma_0)$	$u(\sigma_e)$
$u(\sigma_h)$	

	Valid		Invalid	
Accepted	0	r_e	$-L$	$r_e - L$
	$r_e - c_c$		$-L - c_c$	
Rejected	0	0	0	0
	$-c_c$		$-c_c$	
Pivotal	0	r_e	0	$r_e - L$
	$r_e - c_c$		$-c_c$	

Aggregating payoffs with beliefs

	p_V	p_I
p_A	p_{AV}	p_{AI}
p_R	p_{RV}	p_{RI}
p_P	p_{PV}	p_{PI}

$$\begin{array}{c|c} \mathbb{E}\left(u(\sigma_0)\right) & \mathbb{E}\left(u(\sigma_e)\right) \\ \hline \multicolumn{2}{c}{\mathbb{E}\left(u(\sigma_h)\right)} \end{array}$$

$$\begin{array}{c|c} -p_{AI}\,L & (p_A+p_P)\,r_e-(p_{AI}+p_{PI})\,L \\ \hline \multicolumn{2}{c}{(p_{AV}+p_{PV})\,r_e - c_c - p_{AI}\,L} \end{array}$$

Equations !

$$\mathbb{E}\left(u(\sigma_h)\right) \lesseqgtr \mathbb{E}\left(u(\sigma_0)\right)$$
$$\left(p_{AV} + p_{PV}\right) r_e - c_c - p_{AI} L \lesseqgtr -p_{AI} L$$
$$\left(p_{AV} + p_{PV}\right) r_e \lesseqgtr c_c$$

$$\mathbb{E}\left(u(\sigma_h)\right) \lesseqgtr \mathbb{E}\left(u(\sigma_e)\right)$$
$$\left(p_{AV} + p_{PV}\right) r_e - c_c - p_{AI} L \lesseqgtr \left(p_A + p_P\right) r_e - \left(p_{AI} + p_{PI}\right) L$$
$$p_{PI} L \lesseqgtr \left(p_{AI} + p_{PI}\right) r_e + c_c$$

$$\mathbb{E}\left(u(\sigma_0)\right) \lesseqgtr \mathbb{E}\left(u(\sigma_e)\right)$$
$$-p_{AI} L \lesseqgtr \left(p_A + p_P\right) r_e - \left(p_{AI} + p_{PI}\right) L$$
$$p_{PI} L \lesseqgtr \left(p_A + p_P\right) r_e$$

Payoffs with amendments

	Valid		Invalid	
Accepted	0	r_e	$-L$	$r_e - L - L_e$
	$r_e - c_c$		$-L - c_c$	
Rejected	0	0	0	$-L_e$
	$-c_c$		$-c_c$	
Pivotal	0	r_e	0	$r_e - L - L_e$
	$r_e - c_c$		$-c_c$	

$$\begin{array}{c|c} -p_{AI}\,L & (p_A + p_P)\,r_e - (p_{AI} + p_{PI})\,L - p_I\,L_e \\ \hline \multicolumn{2}{c}{(p_{AV} + p_{PV})\,r_e - c_c - p_{AI}\,L} \end{array}$$

Equations with amendments

$$u(\sigma_h) \lesseqqgtr u(\sigma_0)$$
$$(p_{AV} + p_{PV})\, r_e \lesseqqgtr c_c$$

$$u(\sigma_h) \lesseqqgtr u(\sigma_e)$$
$$p_{PI}\, L + p_I\, L_e \lesseqqgtr (p_{AI} + p_{PI})\, r_e + c_c$$

$$u(\sigma_0) \lesseqqgtr u(\sigma_e)$$
$$p_{PI}\, L + p_I\, L_e \lesseqqgtr (p_A + p_P)\, r_e$$