Fast Plurality Consensus in the Gossip and Population Protocol Model

Gregor Bankhamer Petra Berenbrink Felix Biermeier Robert Elsässer
Hamed Hosseinpour Dominik Kaaser Peter Kling

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

PARIS LODRON UNIVERSITÄT SALZBURG

October 9, 2023

Plurality Consensus

- We consider *plurality consensus* in a distributed system of n agents.
- Initially each agent has one of k possible *opinions*.
- Agents interact in pairs and update their opinions based on other opinions they observe.
- The goal is that eventually all agents agree on the same opinion.
- If there is a sufficiently large *bias* the initially largest opinion should prevail.

- Consensus is a fundamental problem in distributed computing and beyond:
  - fault tolerant sensor arrays
  - majority-based conflict resolution
  - models for dynamic particle systems and biological processes
  - opinion spreading processes in social networks

Undecided State Dynamics

Basic variant: [Angluin et al., Distributed Computing 2008]

- Agents interact in pairs chosen uniformly at random.
- Any agent that encounters another agent with a different opinion becomes *undecided*.
- Undecided agents adopt the first opinion they observe.

Undecided State Dynamics

Basic variant: [Angluin et al., Distributed Computing 2008]

- Agents interact in pairs chosen uniformly at random.
- Any agent that encounters another agent with a different opinion becomes *undecided*.
- Undecided agents adopt the first opinion they observe.

Related work:

- Angluin et al. show that consensus is reached w.h.p. in $\mathrm{O}(n \log n)$ interactions for $k = 2$ opinions.
- Becchetti et al. [SODA'15] analyzes the case $k = \mathrm{O}((n/\log n)^{1/3})$ opinions.
- Condon et al. [Nat. Comput. 2020] reduce the required bias to $\Omega(\sqrt{n \log n})$.
- Clementi et al. [MFCS'18] study the undecided state dynamics in the gossip model.
- They show convergence in $\mathrm{O}(\log n)$ rounds for $k = 2$ opinions, w.h.p.
- Berenbrink et al. [ICALP'16] and Ghaffari and Parter [PODC'16] consider a synchronized variant.
- They achieve consensus in $\mathrm{O}(\log k \log n)$ rounds but require a bias in their analysis.

Undecided State Dynamics

Basic variant: [Angluin et al., Distributed Computing 2008]

- Agents interact in pairs chosen uniformly at random.
- Any agent that encounters another agent with a different opinion becomes *undecided*.
- Undecided agents adopt the first opinion they observe.

We consider two models:

Population Model

- discrete time steps
- one random pair of agents interacts

- number of interactions
- number of states

Gossip Model

- synchronous rounds
- every agent interacts simultaneously

- number of rounds
- memory in bits

Undecided State Dynamics

Basic variant: [Angluin et al., Distributed Computing 2008]

- Agents interact in pairs chosen uniformly at random.
- Any agent that encounters another agent with a different opinion becomes *undecided*.
- Undecided agents adopt the first opinion they observe.

We consider two models:

Population Model

- discrete time steps
- one random pair of agents interacts

- number of interactions
- number of states

Gossip Model

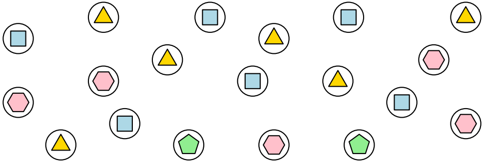
- synchronous rounds
- every agent interacts simultaneously

- number of rounds
- memory in bits

Population Model

[Angluin et al., Distributed Computing 2006]

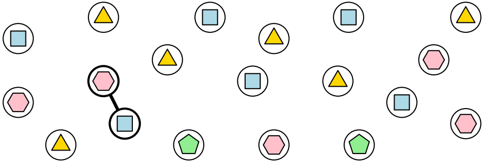
- resource-limited mobile agents (finite-state machines)
- computation is a sequence of pairwise interactions
- interacting agents apply a common transition function δ



Population Model

[Angluin et al., Distributed Computing 2006]

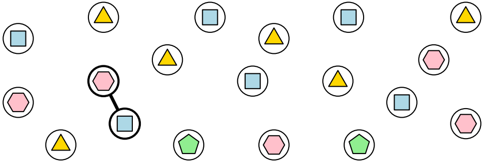
- resource-limited mobile agents (finite-state machines)
- computation is a sequence of pairwise interactions
- interacting agents apply a common transition function δ



Population Model

[Angluin et al., Distributed Computing 2006]

- resource-limited mobile agents (finite-state machines)
- computation is a sequence of pairwise interactions
- interacting agents apply a common transition function δ



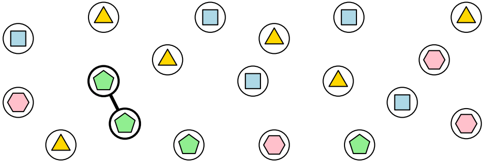
$$\delta : Q \times Q \to Q \times Q$$

$$\delta(⬡, \square) = (⬠, ⬠)$$

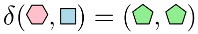
Population Model

[Angluin et al., Distributed Computing 2006]

- resource-limited mobile agents (finite-state machines)
- computation is a sequence of pairwise interactions
- interacting agents apply a common transition function δ



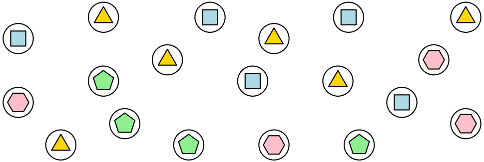
$$\delta : Q \times Q \to Q \times Q$$



Population Model

[Angluin et al., Distributed Computing 2006]

- resource-limited mobile agents (finite-state machines)
- computation is a sequence of pairwise interactions
- interacting agents apply a common transition function δ



Our Contribution

- We consider a *synchronized* variant of the undecided state dynamics.
- A *phase clock* divides time into *phases*.
- Each phase consists of two *parts*.

Actions performed when agents (u, v) interact:

if u is in the decision part:
  if opinion$[u] \neq$ opinion$[v]$ **then do once**
    opinion$[u] \leftarrow$ undecided

if u is in the boosting part:
  if opinion$[u] =$ undecided **then**
    opinion$[u] \leftarrow$ opinion$[v]$

synchronize phase clocks

Our Contribution

- We consider a *synchronized* variant of the undecided state dynamics.
- A *phase clock* divides time into *phases*.
- Each phase consists of two *parts*.

Decision Part

Agents become undecided if they encounter a different opinion.

Actions performed when agents (u, v) interact:

if u is in the decision part:
    if opinion$[u] \neq$ opinion$[v]$ **then do once**
        opinion$[u] \leftarrow$ undecided

if u is in the boosting part:
    if opinion$[u] =$ undecided **then**
        opinion$[u] \leftarrow$ opinion$[v]$

synchronize phase clocks

Our Contribution

- We consider a *synchronized* variant of the undecided state dynamics.
- A *phase clock* divides time into *phases*.
- Each phase consists of two *parts*.

Boosting Part

All undecided agents adopt one of the remaining opinions.

Actions performed when agents (u, v) interact:

if u is in the decision part:
  if opinion$[u] \neq$ opinion$[v]$ **then do once**
    opinion$[u] \leftarrow$ undecided

if u is in the boosting part:
  if opinion$[u] =$ undecided **then**
    opinion$[u] \leftarrow$ opinion$[v]$

synchronize phase clocks

Our Contribution

- We consider a *synchronized* variant of the undecided state dynamics.
- A *phase clock* divides time into *phases*.
- Each phase consists of two *parts*.

Actions performed when agents (u, v) interact:

if u is in the decision part:
  if opinion$[u] \neq$ opinion$[v]$ **then do once**
    opinion$[u] \leftarrow$ undecided

if u is in the boosting part:
  if opinion$[u] =$ undecided **then**
    opinion$[u] \leftarrow$ opinion$[v]$

synchronize phase clocks

- Our protocol reaches consensus in $\mathrm{O}(n \log^2 n)$ interactions.
- If there is a *plurality* opinion, the agents agree on that opinion.
- Otherwise, they agree on a *significant* opinion.

- Our results hold for up to n opinions and independently of a bias.

Analysis

- assume that the phase clocks strictly separate the phases for all agents
- define two series of random vectors $\mathcal{X} = (\boldsymbol{X}(t))_{t \in \mathbb{N}}$ and $\mathcal{Y} = (\boldsymbol{Y}(t))_{t \in \mathbb{N}}$
  - $\boldsymbol{X}_i(t)$: number of agents with opinion i at the beginning of the decision part of phase t.
  - $\boldsymbol{Y}_i(t)$: number of agents with opinion i at the beginning of the boosting part of phase t.

Observation (Decision Part)

Fix $\boldsymbol{X}(t) = \boldsymbol{x}$. Then

$$\boldsymbol{Y}_i(t) \sim \mathrm{Bin}(\boldsymbol{x}_i, \boldsymbol{x}_i/n).$$

Observation (Boosting Part)

Fix $\boldsymbol{Y}(t) = \boldsymbol{y}$ and $d = \|\boldsymbol{y}\|_1$. Then

$$\boldsymbol{X}_i(t+1) \sim \mathrm{PE}(\qquad\qquad).$$

Pólya-Eggenberger Distribution

- The Pólya-Eggenberger process is a simple urn process that runs in discrete steps.
- Initially the urn contains a red balls and b blue balls ($a, b \in \mathbb{N}_0$).
- In each step:
  - draw one ball uniformly at random,
  - observe its color,
  - return the ball, and
  - add one additional ball of the same color.
- The Pólya-Eggenberger distribution is denoted $\mathrm{PE}(a, b, m)$.
- It describes the total number of red balls after m steps.

Analysis

- assume that the phase clocks strictly separate the phases for all agents
- define two series of random vectors $\mathcal{X} = (\boldsymbol{X}(t))_{t\in\mathbb{N}}$ and $\mathcal{Y} = (\boldsymbol{Y}(t))_{t\in\mathbb{N}}$
  - $\boldsymbol{X}_i(t)$: number of agents with opinion i at the beginning of the decision part of phase t.
  - $\boldsymbol{Y}_i(t)$: number of agents with opinion i at the beginning of the boosting part of phase t.

Observation (Decision Part)

Fix $\boldsymbol{X}(t) = \boldsymbol{x}$. *Then*

$$\boldsymbol{Y}_i(t) \sim \mathrm{Bin}(\boldsymbol{x}_i, \boldsymbol{x}_i/n).$$

Observation (Boosting Part)

Fix $\boldsymbol{Y}(t) = \boldsymbol{y}$ *and* $d = \|\boldsymbol{y}\|_1$. *Then*

$$\boldsymbol{X}_i(t+1) \sim \mathrm{PE}(\boldsymbol{y}_i, d - \boldsymbol{y}_i, n - d).$$

Analysis

- assume that the phase clocks strictly separate the phases for all agents
- define two series of random vectors $\mathcal{X} = (\boldsymbol{X}(t))_{t\in\mathbb{N}}$ and $\mathcal{Y} = (\boldsymbol{Y}(t))_{t\in\mathbb{N}}$
  - $\boldsymbol{X}_i(t)$: number of agents with opinion i at the beginning of the decision part of phase t.
  - $\boldsymbol{Y}_i(t)$: number of agents with opinion i at the beginning of the boosting part of phase t.

Observation (Decision Part)

Fix $\boldsymbol{X}(t) = \boldsymbol{x}$. Then

$$\boldsymbol{Y}_i(t) \sim \mathrm{Bin}(\boldsymbol{x}_i, \boldsymbol{x}_i/n).$$

Observation (Boosting Part)

Fix $\boldsymbol{Y}(t) = \boldsymbol{y}$ and $d = \|\boldsymbol{y}\|_1$. Then

$$\boldsymbol{X}_i(t+1) \sim \mathrm{PE}(\boldsymbol{y}_i, d - \boldsymbol{y}_i, n - d).$$

We consider three cases, depending on the number of opinions k.

- Case 1: $k \le \sqrt{n}/\log n$
- Case 2: $\sqrt{n}/\log n < k \le \sqrt{n}$
- Case 3: $\sqrt{n} < k$

Case 1: $k \leq \sqrt{n}/\log n$

- The proof follows along the lines of known results. [Ghaffari and Parter, PODC'16] [Berenbrink et al., ICALP'16]
- Opinions are classified as strong, weak, or super-weak. [Ghaffari and Lengler, PODC'18]
- We consider all pairs of opinions and $\mathrm{O}(\log n)$ phases:
  - at least one opinion in each pair becomes weak, then super-weak, and then extinct.
- For pairs of strong opinions of similar initial size we apply a drift result. [Doerr et al., SPAA'11]

Case 2: $\sqrt{n}/\log n < k \leq \sqrt{n}$

- This case is the main novelty of our analysis.
- It contains many hard configurations:
  - Opinions can be strong and super-weak at the same time.
  - Opinions cannot be tracked via concentration inequalities.
  - Opinions do not vanish immediately.
  - The opinion which provides the maximum support changes over time.

- We consider $\mathrm{O}(\log n)$ phases and exploit the variance of the process.
- There is (at least) one opinion which gains a support of $\Omega(n \cdot \log^{3/2} n)$.
- This follows from the drift result applied to the support of the largest opinion.

- A case distinction and counting arguments yield the following:
- many opinions become small (and eventually die out) in subsequent phases.

- After at most $\mathrm{O}(\log n)$ phases we are back in Case 1.

Case 3: $\sqrt{n} < k$

- It might happen that all agents become undecided.
- In this case, we restore the opinion distribution from the beginning of the phase.
- The probability can be bounded by $(1 - 1/n)^n < 1/e$.
- In all other phases, a constant fraction of the opinions dies out.

- After at most $\mathrm{O}(\log n)$ phases we are back in Case 2.

Our Result

Theorem (simplified)

Our protocol uses $k \cdot \Theta(\log \log n)$ *states per agent.*
All agents agree on a significant opinion in $\mathrm{O}(n \log^2 n)$ *interactions w.h.p.*
If there is an additive bias of order $\omega(\sqrt{n \log n})$*, the initial plurality opinion wins w.h.p.*

Exact Plurality Consensus

Challenges:

- Previous protocol allows opinions with very small support to win with a small probability.
- Any significant opinion wins with good probability.
- Another approach is needed. Approach:

- Allow failure with small probability.
- Keep difference between the largest and any other opinion until the opinion is eliminated.

Exact Plurality Consensus

- Divide the nodes into four different sets: collectors, trackers, players, and clock nodes
- Collectors store opinions - up to 10 of the same kind.
- Trackers keep track of the "tournaments.
- Players perform the tournaments - exact majority among two different opinions.
- Clocks keep track of the time.

Theorem (simplified)

If the opinions are numbered, our protocol uses $\mathrm{O}(k + \log n)$ *states per agent and* $\mathrm{O}(k \log n)$ *parallel time.*
If the opinions are not ordered, our protocol uses $\mathrm{O}(k + \log n)$ *states per agent and* $\mathrm{O}(k \log n + log^2 n)$ *parallel time.*

Conclusions and Open Problems

- Our work's main novelty is the unconditional analysis for any number of opinions and bias.

- One natural open question is whether our results are tight.
- Our algorithm needs $\mathrm{O}(\log n)$ phases for breaking ties.
- It might be possible to work with shorter phase lengths or interleaved consecutive phases.

- For the gossip model it is known that the unsynchronized undecided state dynamics is much slower than the synchronized version.
- It would be interesting to show a similar result for the population model.

- Finally, another open question is to bound the expected running time of the USD.
- Can we design a *stable* protocol that always converges to one opinion with probability 1?

Thank You — Questions welcome!