A semi-topological view of real-world consensus

Galois tech talk

Murdoch J. Gabbay

7 October 2020

Thanks

Thanks to the Galois tech talk organisers for the invitation to speak.

This talk is based on conversations with Giuliano Losa.

What is consensus as a mathematical notion?

In the real world we base decisions on sets of agents we trust or cooperate with: $\{\text{wife}, \text{mother-in-law}\}$, $\{\text{Economist}, \text{BBC}, \text{NYT}\}$, $\{\text{TheWeatherOutlook}, \text{BBC}\}$. These systems of quorums are

- **open** (people and institutions appear and disappear),
- **unpermissioned** ("Do as $\{\text{father}, \text{mother}\}$ say"),
- **not percentage-based** ("Drink beer if >50% of the population does"),
- **local** (no universally accepted central oracle of truth), and
- **mutable** (Henry VIII had six wives, and changed Catholic $\rightarrow$ Church of England).

Question: What is an open permissionless voting system with local mutable quorums, mathematically — and when does it remain organised and coherent?

Observed behaviour is coherent

These systems typically self-organise into (pockets of) stability.

Who we go to lunch with. Who we vote for. What news we believe. What brands we wear. Stellar is an implemented manifestation of such a system. These display stability, and coherence.

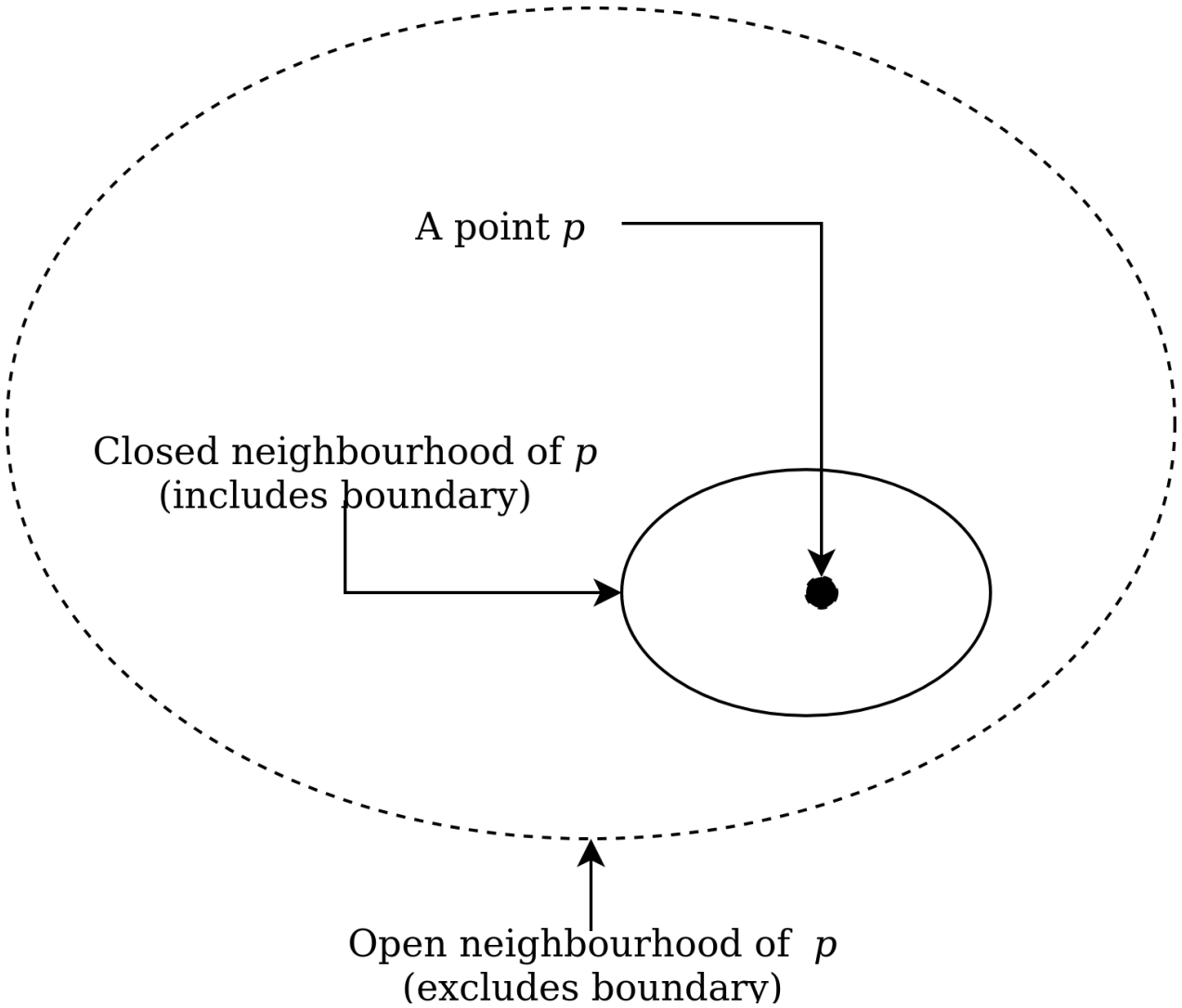
What mathematical properties do they have to explain this?

In this talk I will outline one such mathematics. This is based on conversations with Giuliano, mostly of this paper:

[1] https://losa.fr/research/StellarConsensus/

I'm a mathematician, and I will reach to a familiar tool: topology.

Picture a typical topological space $\mathbb{R}^2$: points and neighbourhoods



Topological space (standard definition)

A topological space $(\mathbb{P}, \mathcal{O}_{\mathbb{P}})$ is:

- a set $\mathbb{P}$ of points and
- a set $\mathcal{O}_{\mathbb{P}} \subseteq pow(\mathbb{P})$ of open sets closed under arbitrary unions and finite intersection:
  1. $\mathcal{P} \subseteq \mathcal{O}_{\mathbb{P}}$ implies $\bigcup \mathcal{P} \in \mathcal{O}_{\mathbb{P}}$ (arbitrary union).
  2. $\mathcal{P} \subseteq_{fin} \mathcal{O}_{\mathbb{P}}$ implies $\bigcap \mathcal{P} \in \mathcal{O}_{\mathbb{P}}$ (finite intersection).

If $p \in P \in \mathcal{O}_{\mathbb{P}}$ then call P a neighbourhood of p.

(A closed set is the complement of an open set. Closed sets are closed under arbitrary intersections and finite unions.)

Semitopological space (novel)

A semitopological space $(\mathbb{P}, \mathcal{O}_{\mathbb{P}})$ is:

- a set $\mathbb{P}$ of points/participants and
- a set of $\mathcal{O}_{\mathbb{P}} \subseteq pow(\mathbb{P})$ of open sets/quorums closed under arbitrary unions.
  1. $\mathcal{P} \subseteq \mathcal{O}_{\mathbb{P}}$ implies $\bigcup \mathcal{P} \in \mathcal{O}_{\mathbb{P}}$
  2. ~~$\mathcal{P} \subseteq_{fin} \in \mathcal{O}_{\mathbb{P}}$ implies $\bigcap \mathcal{P} \in \mathcal{O}_{\mathbb{P}}$~~

If $p \in P \in \mathcal{O}_{\mathbb{P}}$ then call P a neighbourhood or quorum of p.

Called a *semitopology* by analogy with

- *semilattices* (lattices with union but no intersection) or
- *semigroups* (groups without inverses).

Semitopology

The neighbourhoods/quorums of p are intuitively those that may be sufficient to cause p to act.

Semitopologies lack the intersection property since just because P is sufficient for p to act, and also P' is, does not imply $P \cap P'$ is.

Consider quorums of $p = \text{Jamie}$

$$P = \{\text{Jamie}, \text{wife}\} \qquad \text{and} \qquad P' = \{\text{Jamie}, \text{mother-in-law}\}.$$

Either set is sufficient to organise a lunch out, but not $P \cap P' = \{\text{Jamie}\}$.

Let's use semitopologies to explore notions of consensus.

Continuity = consensus

Call $f : \mathbb{P} \to X$ continuous at $p \in \mathbb{P}$ when f is constant on some neighbourhood of p. In symbols:

$$\exists P \in \mathcal{O}_{\mathbb{P}}. \Big(p \in P \ \wedge \ \forall p' \in P. f(p') = f(p) \Big).$$

- If f is continuous at p, then intuitively p *agrees with a quorum* (wrt f).
- If f is discontinuous at p then intuitively p *disagrees with all its quorums* (wrt f).

(For experts: we gave X the discrete topology.)

Fix some set F of 'admissible' or 'observable' f:

- If all $f \in F$ are continuous at p, call p well-behaved (wrt F).
- If some $f \in F$ is discontinuous at p, call p Byzantine (wrt F).

Continuity = consensus

Example: Participants $p : \mathbb{R}$ and observables are $f : \mathbb{R} \to \mathbb{R}$ that are continuous on $\mathbb{R} \setminus \{0\}$, e.g.

$$f(x) = \text{if } x \leq 0 \text{ then } 0 \text{ else } 1.$$

Then $0 \in \mathbb{R}$ is the only Byzantine participant, since if $x_i \to 0$ from below then $f(x_i) \to 0$, but if $x_i \to 0$ from above then $f(x_i) \to 1$.

Example: We can *a priori* fix a set $P \subseteq \mathbb{P}$ of well-behaved participants (typically this is done in the field), then call f observable precisely when f is continuous on P. Unsurprisingly, P becomes the set of well-behaved participants and anything else becomes Byzantine.

Problem of consensus = compute an observable $f : \mathbb{P} \to X$.

Refinements: blocking sets and partiality

Blocking sets: Consider $f : \mathbb{R} \to \mathbb{R}$ continuous away from $\{1,\ 1/2,\ 1/3,\ 1/4,\ \dots\}$.

Then 0 is a limit point of Byzantine participants. What does 'well-behaved' mean, if you're surrounded by crazies?

In the Stellar literature this is called befouled, and $\{1, 1/2, 1/3, 1/4, \dots\}$ is called a blocking set for 0.

Topologically, this is a set with a limit point. (A limit point $p \in \mathbb{P}$ of $B \subseteq P$ is such that $B \cap P \neq \varnothing$ for every neighbourhood $p \in P \in \mathcal{O}_{\mathbb{P}}$.)

B blocking set for $p = p$ limit point for B.

Refinements: blocking sets and partiality

Partiality: Malicious actors may withhold a return value.

Consider

$$f(x) = 1/x.$$

Here f isn't just discontinous at 0; it's undefined.

That's fine: take $f : \mathbb{P} \rightharpoonup X$ partial (or add a $\perp$ value), and on well-behaved participants f is *defined* (or f returns $\perp$) and continuous.

For me, semitopologies provide a useful language with which to express and explore such refinements. If you have further ideas, please be in touch.

Consensus cluster

Suppose $(\mathbb{P}, \mathcal{O}_{\mathbb{P}})$ is a semitopological space.

Notation: If $X, Y \subseteq \mathbb{P}$ then write

$$X \vdash Y \quad \text{when} \quad X \cap Y \neq \varnothing.$$

(V. simple judgement. Note that $X \vdash Y \Leftrightarrow Y \vdash X$.)

Definition: Call $C \subseteq \mathbb{P}$ a cut set or consensus cluster when

$$\forall P, P' \in \mathcal{O}_{\mathbb{P}}. \quad P \vdash C \wedge C \vdash P' \quad \Rightarrow \quad P \vdash P'.$$

Reformulation: C is a cut set when for any two $c, c' \in C$ and neighbourhoods $c \in P$ and $c' \in P'$, P and P' must intersect.

I haven't seen this condition in the topology literature. It is a form of strong converse to the Hausdorff property:

Cut set and continuity

Definition: Call $C \subseteq \mathbb{P}$ a cut set or consensus cluster when

$$\forall P, P' \in \mathcal{O}_{\mathbb{P}}.\ \ P \vdash C \wedge C \vdash P' \ \ \Rightarrow \ \ P \vdash P'.$$

Lemma: Suppose $f : \mathbb{P} \to X$ and $C \subseteq \mathbb{P}$ is an open cut set. Then if f is continuous on C, then f is constant on C:

$$\forall c, c' \in C. f(c) = f(c').$$

Intuitively: C has consensus about the value of f.

Proof: Since f is continuous, it must be constant on two open neighbourhoods $c \in P$ and $c' \in P'$. These intersect, so $f(c) = f(c')$.

Lemma 4

Definition: Call $C \subseteq \mathbb{P}$ a cut set or consensus cluster when

$$\forall P, P' \in \mathcal{O}_{\mathbb{P}}.\ \ P \vdash C \wedge C \vdash P' \ \ \Rightarrow \ \ P \vdash P'.$$

Proposition ([2, Lemma 4]): If C and C' are intersecting open cut sets then so is $C \cup C'$.

Proof: Suppose $P \vdash C \cup C'$ and $C \vdash C'$ and $C \cup C' \vdash P'$. Without loss of generality suppose this is because $P \vdash C$ and $C' \vdash P'$ (other cases are no harder).

Then $P \vdash C \vdash C' \vdash P'$ and (since C and C' are open) also $P \vdash P'$.

Corollaries

Proposition ([2, Lemma 4]): If C and C' are open cut sets then so is $C \cup C'$.

Corollary: If $C_1 \subseteq C_2 \subseteq \ldots$ is an ascending chain of open cut sets then $\bigcup_i C_i$ is an open cut set.

Further corollary: $(\mathbb{P}, \mathcal{O}_{\mathbb{P}})$ partitions itself into:

- a collection of disjoint maximal cuts (maximal consensus clusters), along with
- some other points that are not in any cut set.

These cut sets are areas of consensus. This goes some way to explaining Stellar's stability: the system naturally partitions itself into maximal open cut sets and, given sufficient connectivity, there is likely to be only one such.

Quorum = open set

A key property in [1] is:

Property 1 (Quorum sharing). *If Q_p is a quorum of p and $p' \in Q_p$ then there exists a quorum $Q_{p'}$ of p' such that $Q_{p'} \subseteq Q_p$.*

In [1] each point has its own set of quorums.

A semitopology abstracts this such that an open set is a quorum for every point it contains.

After all: if P agrees and $P' \subseteq P$, then P' agrees; so if P' can cause p to act then so can P.

[1, Property 1] in the semitopological view becomes:
quorum = open set.

Why "Semi-"?

Semigroups are 'just' groups without inverses (or unit). But semigroup theory is a distinct field with its own character.

Semilattices are 'just' lattices with join but no meet (or vice versa). Likewise, semilattice theory is a distinct field.

Similar definitions can yield rather distinct bodies of theory, concerned with distinct classes of models, and theorems about them.

It's quite instructive to look at the 'median' models of topologies and semitopologies:

The median (semi)topology is . . .

The ‘median’ topology is $\mathbb{R}$, and the ‘median’ function of interest on $\mathbb{R}$ is continuous maps from $\mathbb{R}$ to $\mathbb{R}$.

In topology, broadly speaking points are infinite, we focus on *separation* axioms (I count 18 on this Wikipedia page), and functions are continuous.

In semitopology, at least applied to consensus,

- points tend to be *finite*,
- we focus on *nonseparated* (*intertwined*) spaces, and
- functions are mostly continuous — *but* it’s key that they might be discontinuous on Byzantine participants.

Multiple maximal cut sets (consensus clusters) are to be avoided; we want one consensus cluster, for a single consensus on truth.

Summary: semitopological dictionary

Personal Byzantine Quorum System	$\approx$	semitopological space
quorums	$\approx$	open sets
consensus	$\approx$	continuity
consensus cluster	$\approx$	cut set
max. consensus clusters	$\approx$	partition into max. cuts
[2, Lemma 4]	$\approx$	transitivity of cut
B blocking set of p	$\approx$	p limit point of B
And more:		
slices	$\approx$	closed sets

Computing open sets

Suppose that

- we have $S \subseteq \mathbb{P}$ and
- we want to compute an open neighbourhood $S \subseteq O(S) \in \mathcal{O}_{\mathbb{P}}$.

In practice, S may be participants that are trusted or indispensable to *us* — but to join the system we need a *quorum*/*open set* that includes our trusted S (cf. Slide 17).

Since an open is a quorum for every point it contains, we may need to expand S with more elements for the people who the people we trust, trust, and so forth.

Is there a practical, efficient computation to do this, and how much of $(\mathbb{P}, \mathcal{O}_{\mathbb{P}})$ does it need to build / query?

Computing open sets

Suppose we have $S \subseteq \mathbb{P}$ and wish to compute an open neighbourhood $S \subseteq O(S) \in \mathcal{O}_{\mathbb{P}}$.

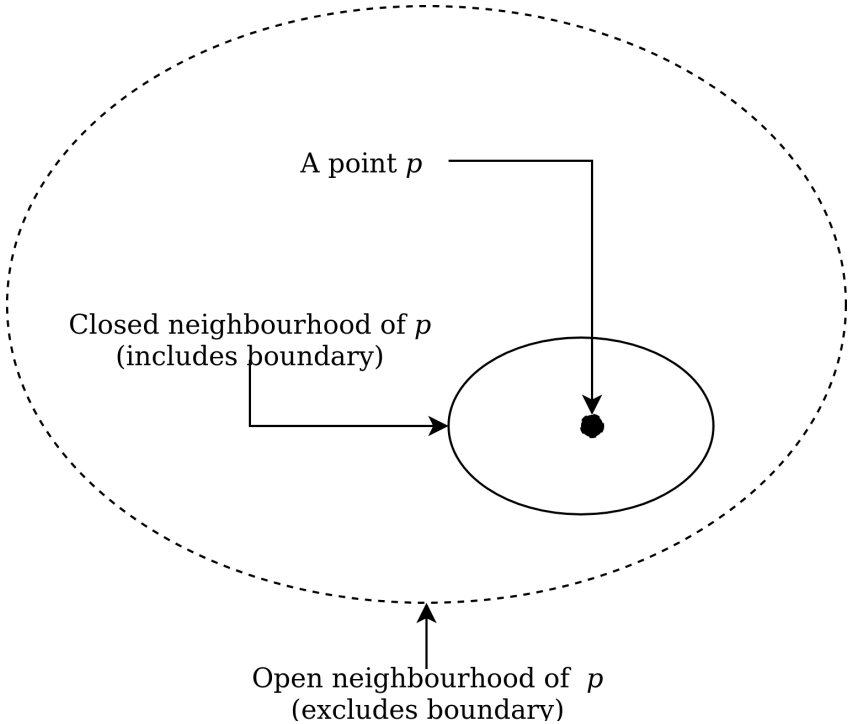
- We could search all opens for one containing S.
- For each $s \in S$ we could choose some open $s \in O(s) \in \mathcal{O}_{\mathbb{P}}$ and take $O(S) = \bigcup_{s \in S} O(s)$.

But this is inefficient, and worse, it may be circular: it replaces calculating *one* $O(S) \in \mathcal{O}_{\mathbb{P}}$ with computing *many* open sets.

We may not have access to this information.

Is there another way? Let's recall our diagram:

A typical (semi)topological space



1. $P = \{P' \in \mathcal{O}_{\mathbb{P}} \mid P' \subseteq P\}$ (P union of open neighbourhoods).
2. $P = \{C \in \mathcal{C}_{\mathbb{P}} \mid C \subseteq P\}$ (P union of closed neighbourhoods).

Computing open sets

In Stellar, a point does not nominate its opens; it nominates a selection of slices.

To me, slices look like *closed* neighbourhoods. A slice can have a *boundary*; that is, it need not have a neighbourhood / quorum for each of its elements.

Let's set: slice = closed neighbourhood.

If we have (at least) one slice for each point, opens can be computed on demand by a fixedpoint: Given e.g. $S \subseteq \mathbb{P}$,

- iteratively add a slice of each point as required until
- the result contains a closed neighbourhood of every point.
- Then stop.

This is taken to be an open set / quorum.

Slice = closed neighbourhood

Note: A slice is *not* necessarily the complement of a quorum, and slices are not closed under intersections.

But these are semitopologies; we shouldn't expect everything to be the same as topologies.

What we do have is that slices are like opens but may have a boundary, and every open is the union of closed neighbourhoods of its points. This turns out to be computationally useful.

So we have: slices = closed neighbourhoods.

Future work (near view)

No immediate algorithmic implications (but see next slide).

Topology is a compelling diagrammatic language. Helpful for communication & useful in outreach, exposition, and perhaps for the statements and verifications of algorithms and their properties.

This links consensus problems with mathematical structures and their intuitions in a new way; algorithms fit into a broader landscape and language; and proofs may have seemed long and arbitrary become short and natural.

We gain a new perspective on what's going on, which may contribute to the system becoming easier to motivate.

Future work (speculative view)

Great if this gives topologists a new path to blockchain research.

To me, this feels like topology + combinatorics and computation.

E.g. we want a semitopological space to consist of a single cut set: to me this feels like connectivity graphs, with (semi)topological structure.

A literature exists on graph connectivity; conditions under which a randomly chosen graph is likely to be fully connected; and algorithms for computing well-behaved subgraphs of graphs. Perhaps maths might be usefully imported from these fields.

Thank you for listening.