The semitopology of heterogeneous consensus

Murdoch J. Gabbay (joint work with Giuliano Losa)

Online Social Choice and Welfare seminar

1 November 2022

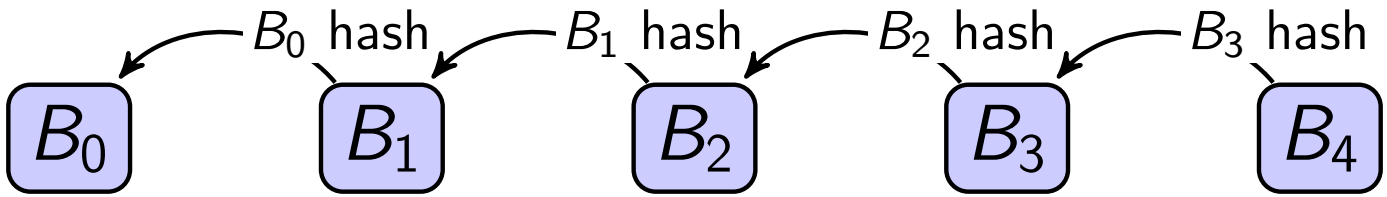
This talk reports on a mathematical analysis of a blockchain-based payments system called Stellar. The maths occupies the body of this talk . . . but let's motivate it first:

Part 0:
What's a blockchain?

A (brief) explanation of blockchain

A blockchain is a distributed database in which each state update B_{t+1}, called a block, is signed with a cryptographic hash of the preceding block B_t.

So: a *blockchain* is a block-chain in which each block cryptographically points to its predecessor:



Say we're at block B_4 at time t=4. How do we progress the system to create and agree on a suitable next block B_5?

Users propose transactions (e.g. *transfer n tokens from X to Y*) to a queue. Validators validate and parcel these up into candidates B_5, B_5', B_5'' at e.g. 1500 transactions per candidate.

Validators then vote on which candidate B_5 to add. It's this voting procedure that interests us.

Consensus must be efficient, permissionless, . . .

This voting problem has some specific requirements:

- A new block needs chosen every few seconds.
  So efficiency counts, this being a worldwide network with limited bandwidth, and a design priority is to avoid recounts or multiple voting rounds, since this costs time and bandwidth!
- The system is (usually) permissionless: new validators can join and old validators can leave.
- It's a consensus problem: that *some* agreement is reached on the next block of transactions, matters more than which particular block is chosen, but also . . .
- . . . the system needs to be resistant to participants trying to manipulate or subvert the system, because . . .

. . . and consensus must be robust

. . . transactions may carry value (e.g. payments, contracts). Hostile behaviour includes:

- DoS (denial-of-service) the consensus algorithm! E.g. vandalism, ransom, infrastructure DoS, or driving users to a competing system.
- Selectively delay consensus for certain blocks; e.g. for front-running, arbitrage, disadvantaging competitors, etc.
- Manipulate blocks while keeping them valid, e.g. to split (fork) the system, double-spend tokens, reverse signed contracts, etc.

(In practice, Bitcoin and Ethereum are highly adversarial environments. This talk will not be about concrete attacks on blockchains; that's another talk which I'm happy to deliver on request.)

How to handle this?

Who votes, and what coalitions can progress?

This is a deep and subtle question, but at a high level there are just two questions:

1. How are votes allocated?
2. What are the progressing coalitions = sets of participants with voting power to progress the system?[1]

Possible design decisions of real-world blockchain systems fall into various categories. We consider three:

[1] *I made up the term 'progressing coalition'. The concept is like a 'winning coalition' in social choice theory, but there's a subtlety: the system is distributed, asynchronous, and permissionless — so state updates must start local and be asynchronously propagated (or not!). I therefore write 'progress' instead of 'win'.*

Who votes, and what's a progressing coalition?

($\propto$ = proportional to)

- Proof-of-Work (PoW):
  Votes $\propto$ compute; progressing coalition = majority vote.
  Unecological since compute = energy = CO_2.
- Proof-of-Stake (PoS):
  Votes $\propto$ stake (tokens/wealth); progressing coalition = majority vote.
  Explicitly equates wealth with governance power.[2]
- Proof-of-Agreement (PoA):
  Votes $\propto$ reputation; progressing coalition = open neighbourhood.
  This is Stellar's approach and it motivates semitopologies.

[2] An interesting wrinkle with PoS is that deciding who has stake is itself a consensus problem. Technical solutions can get a bit baroque.

Comments on the politics of consensus

With PoW and PoS, progressing coalitions are automatically determined from compute and stake respectively. Nobody has to like you or know who you are for you to be in their progressing coalition.

This encodes a libertarian governance structure which is amoral and areputational.[3]

With PoA, progressing coalitions are explicitly specified, and admission is by explicit peer invitation (as we shall see). Thus reputation and trust — social constructs — are represented within the system, distinct from computational power and wealth.

This creates incentives to curate reputation and to play not only *according to the rules* but also according to peer's *moral expectations* — on pain of being removed from their progressing coalitions. *E.g. a front-running PoA validator may be exposed to scandal and excluded from progressing coalitions (front-running PoW and PoS validators can and do operate).*

[3] I made this word up.

Part 0: What's a blockchain?

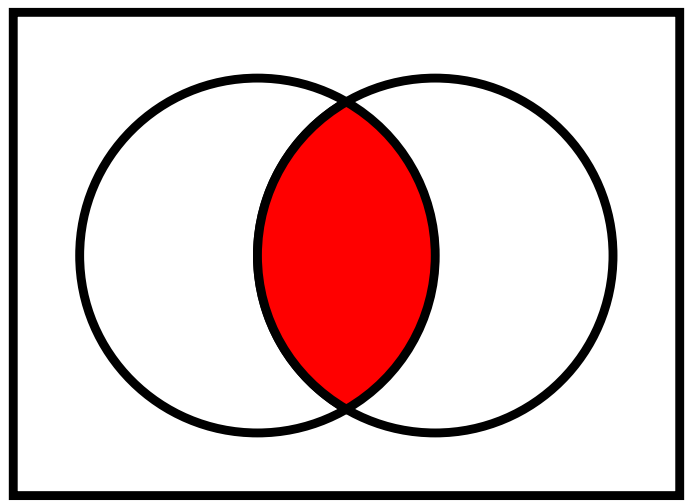
Part 1:
Semitopology, continuity, topens

Definition of a semitopology

Definition. A semitopology is a pair $(P, \mathrm{Open} \subseteq pow(P))$ of

- P a nonempty set of points and
- Open a set of open sets such that $P \in \mathrm{Open}$ and $\mathcal{O} \subseteq \mathrm{Open} \Rightarrow \bigcup \mathcal{O} \in \mathrm{Open}$.

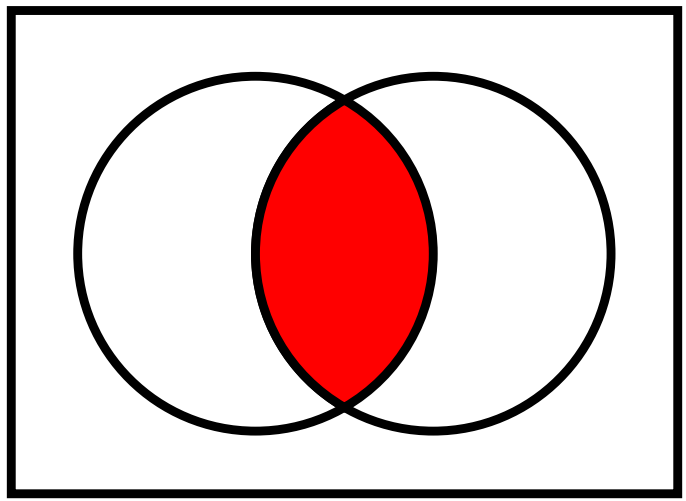
Think: *"topology, minus condition that $\cap$ of two opens is open".*



(Image credit: Wikipedia.)

Key difference from topologies

In a topology, a *minimal* open neighbourhood of p is also *least* (any two minimal open neighbourhoods of p can be intersected, which by minimality is equal to both) . . .



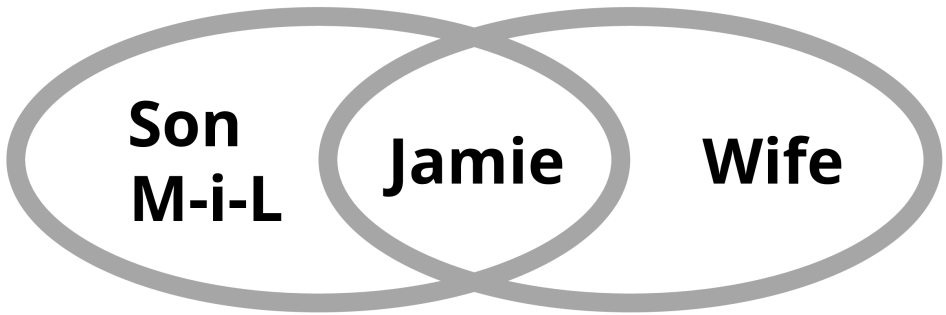
. . . whereas in a semitopology, p may have multiple minimal open neighbourhoods.

Semitopologies have a rich mathematical structure — there is much more to them than being 'weak topologies', much as semigroups are much more than 'weak groups'.

Relevance to consensus

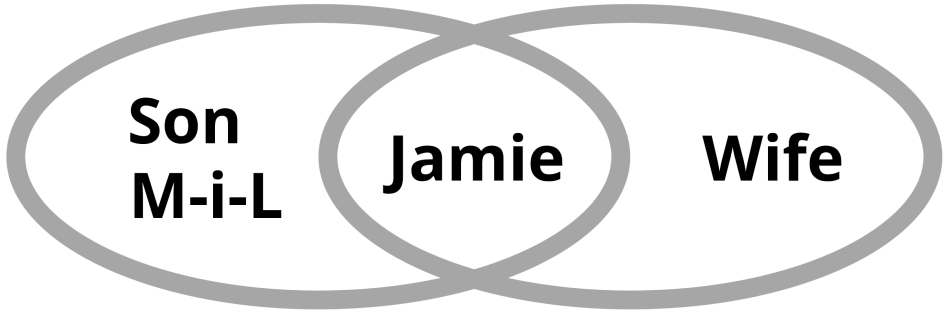
- Interpret $p \in \mathrm{P}$ as a *participant*.
- Interpret an open neighbourhood $p \in O \in \mathrm{Open}$ as a *progressing coalition*.

Worked example: Consider four participants, in two (least) opens (M-i-L = Mother-in-Law):



We seek to progress on whether to wear a smart shirt ('S-shirt') or a T-shirt ('T-shirt') for this lecture. If at least one of {**Jamie**, **Wife**} and {**Jamie**, **Son**, **M-i-L**} agree, then we can progress.

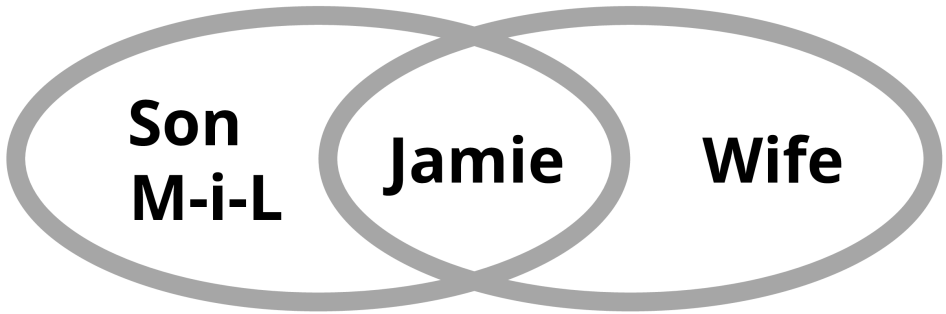
Relevance to consensus



Let's consider the possibilities; list exhaustive modulo permuting between 'S-shirt' and 'T-shirt':

- **Son** and **M-i-L** say 'S-shirt'; **Wife** says 'S-shirt'.
  Jamie must vote 'S-shirt' to progress.
- **Son** and **M-i-L** say 'T-shirt'; **Wife** says 'S-shirt'.
  Jamie can progress with 'S-shirt' or 'T-shirt'.
- **Son** says 'T-shirt', **M-i-L** says 'S-shirt'; **Wife** says 'S-shirt'.
  Jamie can only pregress with 'S-shirt' (or convince **Son** or **M-i-L** to change vote).
- **Son** says 'S-shirt', **M-i-L** says 'T-shirt'; **Wife** says 'S-shirt'.
  (As previous case.)

Relevance to consensus



Note:

- In this example, an easy way for **Jamie** to progress is to agree with **Wife**. That's a special property of a two-element progressing coalition. *(Just goes to show: maths is the best way to understand relationships.)*
- Opens need not be uniform, e.g. my mother-in-law may have her own opens (not illustrated above).

Semitopologies are not obviously self-organising

Consider an arbitrary semitopology. As mentioned, the notion of progress is local and asynchronous, and there are no restrictions on how opens are formed.

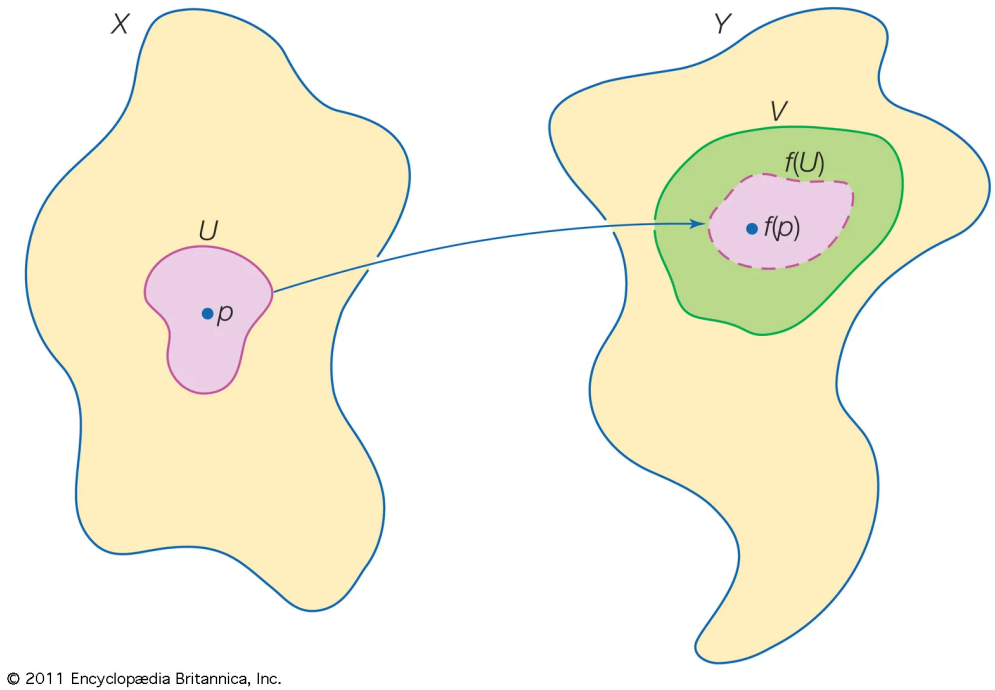
One might expect this to block or dissolve into chaos, but empirically it's stable. It turns out there are solid mathematical reasons for this ...

Continuity (textbook def)

Suppose $(\mathrm{P}, \mathrm{Open})$ and $(\mathrm{P}', \mathrm{Open}')$ are semitopologies. Call a function $f : \mathrm{P} \to \mathrm{P}'$ continous at $p \in \mathrm{P}$ when

- $\forall$ open neighbourhood $f(p) \in O'$,
- $\exists$ open neighbourhood $p \in O \subseteq f^{-1}(O')$.

$$\forall O' \in \mathrm{Open}'. f(p) \in O' \Rightarrow \exists O \in \mathrm{Open}. p \in O \subseteq f^{-1}(O').$$



(Image credit: Britannica.)

Continuous value assignments

Definition.

- Fix a set of values Val with the discrete semitopology $(\mathrm{Val}, pow(\mathrm{Val}))$ in which $\{v\}$ is open for every $v \in \mathrm{Val}$.
- Call a function $f : \mathrm{P} \to \mathrm{Val}$ a value assignment.

Lemma. The following are equivalent:

- f continuous at p.
- $f^{-1}(f(p)) \in \mathrm{Open}$.
- An open agrees with p on its value $f(p) \in \mathrm{Val}$.

This notion of consensus is an instance of topological continuity.

This links consensus to topology (1895, Poincaré) and we get a slogan:

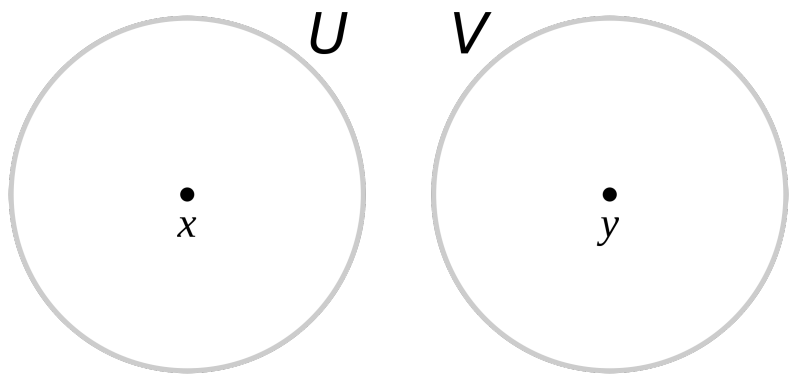
Consensus = Continuity.

Something familiar: Hausdorff separation

Notation. Write $O \between O'$ when $O \cap O' \neq \varnothing$.

You may be familiar with the (standard) Hausdorff property that $p \neq p' \in \mathrm{P}$ have disjoint open neighbourhoods:

$$\exists O, O' \in \mathrm{Open}.(p \in O \wedge p' \in O') \wedge \neg(O \between O').$$



(Image credit: Wikipedia.)

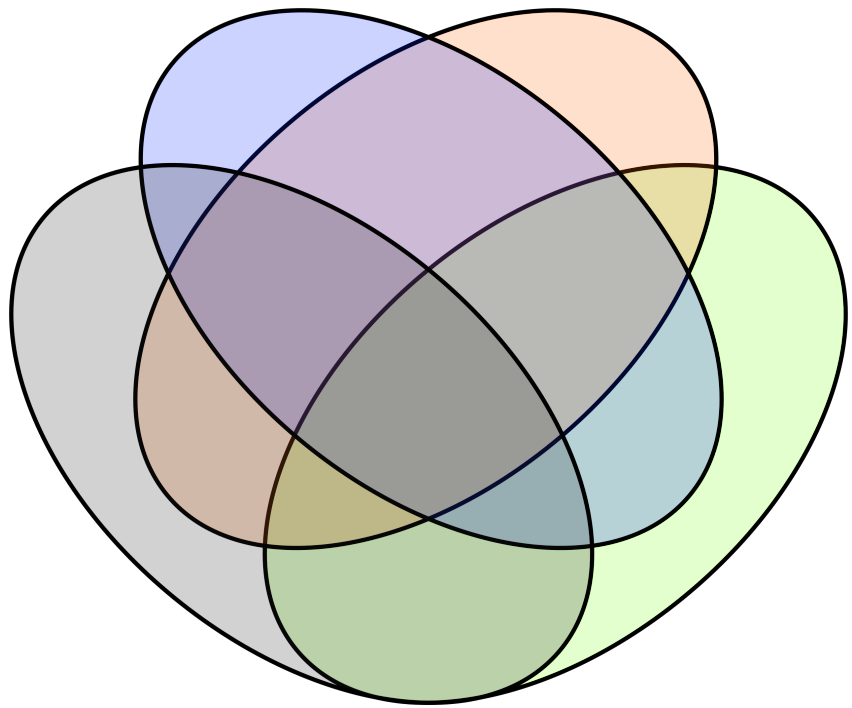
We introduce a novel *anti-Hausdorff* property:

Intertwined points: the anti-Hausdorff property

Definition. Call p and p' intertwined and write $p \between p'$ when

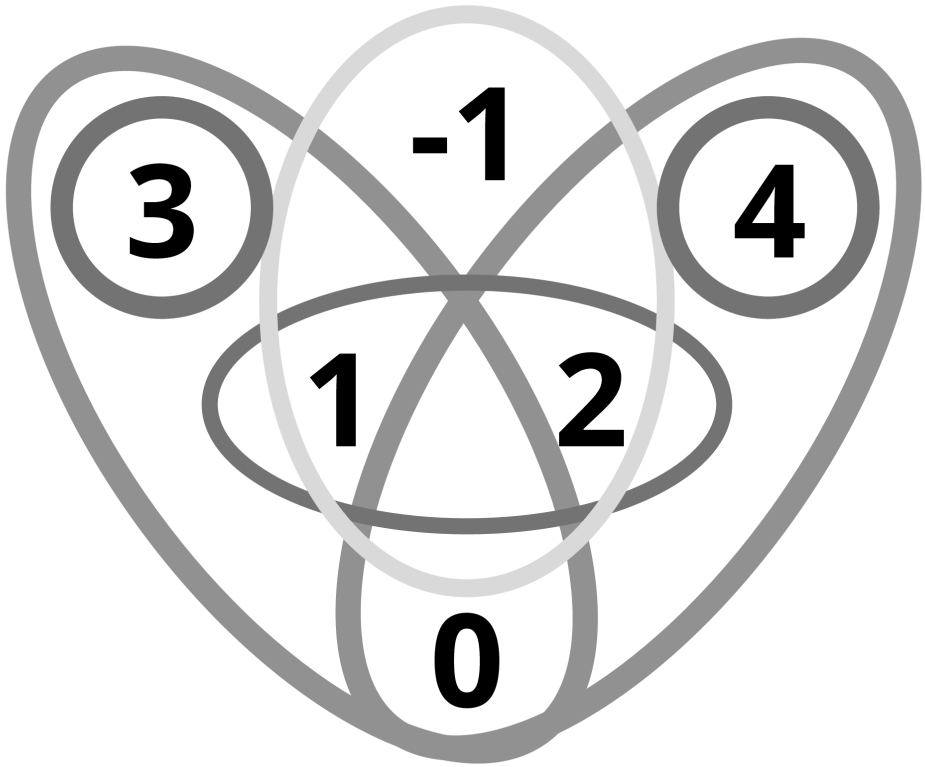
$$\forall O, O' \in \mathrm{Open}.(p \in O \wedge p' \in O') \Rightarrow O \between O'.$$

So $p \between p'$ when all their open neighbourhoods intersect, i.e. the very opposite of Hausdorff separation:



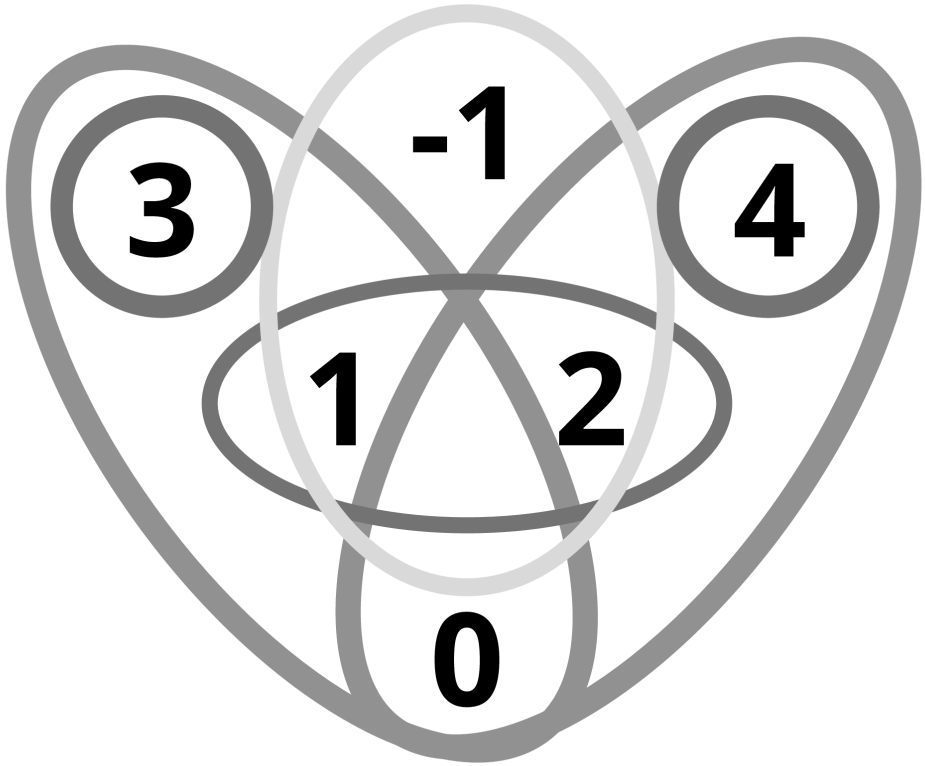
(Image credit: Wikipedia.)

Intertwined points: the anti-Hausdorff property



Q. Which points here are intertwined?

Intertwined points: the anti-Hausdorff property



Q. Which points here are intertwined?

A. -1 ⟩⟨ 0 ⟩⟨ 1 ⟩⟨ 2

Consensus = continuity

Recall: *values* Val with the discrete semitopology in which $\{v\}$ is open for every $v \in \mathrm{Val}$. A *value assignment* is a function $f : \mathrm{P} \to \mathrm{Val}$, and $f : \mathrm{P} \to \mathrm{Val}$ is *continuous* at $p \in \mathrm{P}$ when $f^{-1}(f(p)) \in \mathrm{Open}$.

Lemma. If $p \between p'$ then $f(p) = f(p')$.

Proof. $f^{-1}(f(p)) \ni p$ and $f^{-1}(f(p')) \ni p'$ are open (by continuity), and so (since $p \between p'$) they intersect.

Corollary 1. Continuous value assignments *agree* between intertwined points.

Transitive sets

Definition. Call $S \subseteq \mathrm{P}$ transitive when

$$\forall O, O' \in \mathrm{Open}. O \between S \between O' \Rightarrow O \between O'.$$

Call S topen when it is transitive and open.

Lemma 1. S is topen $\Leftrightarrow$ S is open and $\forall p, p' \in S. p \between p'$ (in words: its points are pairwise intertwined).

Proof. Suppose S is topen and $p, p' \in S$ and $p \in O$ and $p' \in O'$. Then $O \between S \between O'$ and so $O \between O'$.

Conversely, if all points are pairwise intertwined and $O \between S \between O'$ then $p \in O$ and $p' \in O'$ for $p, p' \in S$ and so $O \between O'$.

Topens are interesting because topen S *must agree* (by Lemma 1 and Corollary 1) and *can progress* (since it's open).

Transitive sets

Lemma 2. $\mathcal{S}$ is a set of pairwise intersecting topens $\Rightarrow \bigcup \mathcal{S}$ is topen.

Proof. $O \between \bigcup \mathcal{S} \between O'$ implies (wlog) $O \between S \between S' \between O'$ for some $S, S' \in \mathcal{S}$, and by transitivity $O \between O'$.

Theorem 1 (self-organisation).

1. (P, Open) partitions into disjoint maximal topen sets (plus isolated points).
2. $f : \mathrm{P} \to \mathrm{Val}$ is constant on each partition, where it is continuous.

Proof. By Lemma 2, if topen S and S' intersect then $S \cup S'$ is topen. Also using Lemma 2, an increasing chain $S_0 \subseteq S_1 \subseteq \ldots$ of topens is topen. The partitioning follows. Continuous value assignments are constant from Lemma 1 (all points intertwined) and Corollary 1 (value assignment is constant on intertwined points).

Stellar computes continuous value assignments

Theorem 1 provides a high-level account of consensus in Stellar:

1. Let participants choose freely whom they trust.
2. Derive a semitopology from these local choices by taking suitable unions (using *witness functions*; that's another talk).
3. Compute continuous value assignments.

Theorem 1 proves a semitopology must self-organise into topen communities of local consensus. Note that this resembles how in real life, people self-organise into communities with shared values (+ outliers).

Part 0: What's a blockchain?

Part 1: Semitopology, continuity, topens

Part 2:
Community, kernel, dictator sets

The community of a point

- Call p regular when $p \in T$ for some topen T (open of intertwined points).
- If p is regular then by Theorem 1 it is contained in some maximal topen.
- Call this maximal topen the community of p, where this exists, and write it $K(p)$.

By Theorem 1, if $f : \mathrm{P} \to \mathrm{Val}$ is continuous then

$$f(p) = f(K(p)).$$

In words: *under continuous value assignments, a regular p agrees with its community.*

The kernel of a regular p

Definition. An atom is a minimal nonempty open set. Write $\varnothing \lessdot A$ When $A \in \mathrm{Open}$ is an atom.

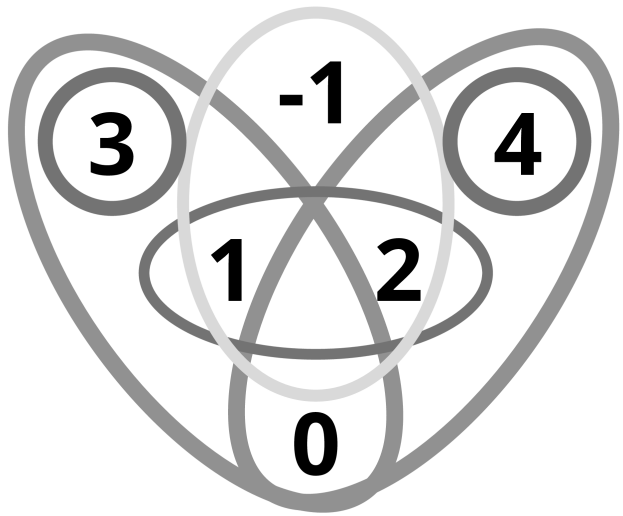
Definition. Suppose p is regular, so $p \in K(p) \in \mathrm{Open}$. Define $ker(p)$ the kernel of p to be the union of the atoms in $K(p)$. In symbols:

$$ker(p) = \bigcup \{A \mid \varnothing \lessdot A \subseteq K(p)\}.$$

The kernel of p is the *union of those minimal opens guaranteed to intersect any open within p's community*.

The kernel of a regular p

$$ker(p) = \bigcup \{A \mid \varnothing \lessdot A \subseteq K(p)\}.$$



- -1 ≬ 0 ≬ 1 ≬ 2
- $ker(\text{-}1) = ker(0) = ker(1) = ker(2) = \{\text{-}1, 1, 2\}$
- $ker(3) = \{3\}$ and $ker(4) = \{4\}$

The kernel theorem (shades of Arrow's theorem)

Theorem 2 (dominance of kernels). Suppose that:

- $f : \mathrm{P} \to \mathrm{Val}$ is a value assignment.
- $p \in \mathrm{P}$ is regular (so $p \in K(p)$).
- f is continuous at p and on some kernel atom $\varnothing \lessdot A \subseteq ker(p)$.

Then $f(p) = f(A)$.

Proof. By assumption $f^{-1}(f(p)) \ni p$ is open. $K(p)$ is intertwined and $A \subseteq K(p)$, so $f^{-1}(f(p)) \between A$. By Theorem 1, $f(A) = f(p)$.

In words: a regular point p is dominated by any of its kernel atoms.

The kernel is a kind of dictator set. It doesn't set out to be this, and its points need not necessarily even know they are in a kernel. This just emerges from the mathematics of semitopologies.

Part 0: What's a blockchain?

Part 1: Semitopology, continuity, topens

Part 2: Community, kernel, dictator sets

Part 3: Conclusion

Concluding remarks

With PoW and PoS, admission into progressing coalitions is determined by compute and stake respectively. Nobody has to like you or know who you are. This encodes an amoral, areputational, libertarian governance structure.

With PoA, admission into progressing coalitions is explicit by peer invitation. Reputation and trust are directly represented within the system, distinct from computational power or wealth.

This creates incentives to curate reputation, and to behave not only according to the rules but also according to peer's moral expectations.

Concluding remarks

PoA is closer to how governance works in real life. Indeed, real blockchains (even PoW/PoS ones) make their substantive governance decisions based on trust and reputation!

It's just not encoded in the system.

Ethereum provides two famous examples:

- When Ethereum forked (into ETC and ETH) after the DAO hack, this was a social decision — by which I mean that the society of Ethereum users debated and made a choice to fork the blockchain. It was not decided within the chain itself.
- Similarly when Ethereum completed the Merge (switching from PoW to PoS), again this was a social decision.

Concluding remarks

There is a breed of blockchain-based distributed systems distinguished by being

- *permissionless* (participants can freely leave and join),
- *heterogeneous* (distributed, asynchronous, diverse),
- *algorithmic* (practical on available hardware and networks), and
- *robust* (against hostile participants, network outage, etc),
- (for Stellar) *dynamic* (you can choose and change whom you trust, based on observed behaviour).

Solutions to these constraints are far from value-neutral: they reflect social views.

Solutions to these constratins are not purely academic: they are responding to business imperatives in real time (e.g. security, performance, competitiveness, and profitability), while seeking to remain true to social ideals (e.g. fairness and decentralisation).

Concluding remarks

I propose there may be much to say about the design, fair governance, and mathematics of this new breed of distributed systems. For example:

1. Can a rigorous connection be made between Theorem 2 and Arrow's theorem?
2. Computational applications of semitopologies to design efficient reliable and robust blockchain algorithms.
3. Study of semitopology-based governance systems, especially for the permissionless distributed asynchronous setting characteristic of blockchains.
4. What are suitable and useful notions of path and homotopy in semitopologies?

Anyone up for research into Semitopological Social Choice?

Bonus slide: Untrusted forks are automatically ignored

Theorem 1 does not guarantee that the semitopology consists of a single topen; Indeed, a semitopology may partition into multiple topens — this is a feature, not a bug!

Consider a topen T that is under attack. Yes the attacker can create a topen J of junky opens, but J remains disjoint from T, so the attack fails, having no effect on T.

With PoW or PoS systems, an attacker *just has to offer work or stake* to be able to (try to) destabilise existing actors.

Theorem 1 gives Stellar a surprisingly robust assurance of stability: an attacker must convince participants to explicitly trust it, by adding it to their opens. It's a different approach.