

The (Normative) Case for Bridging-Based Ranking

Luke Thorburn

March 2024

Context

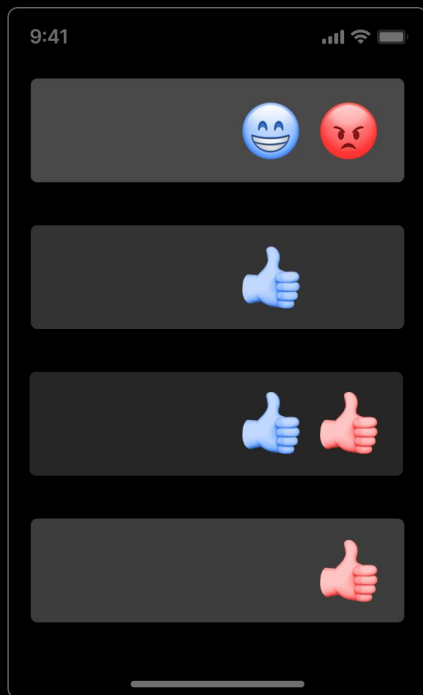
bridging An increase in mutual understanding and trust across divides, creating space for productive conflict, deliberation, or cooperation.

Context

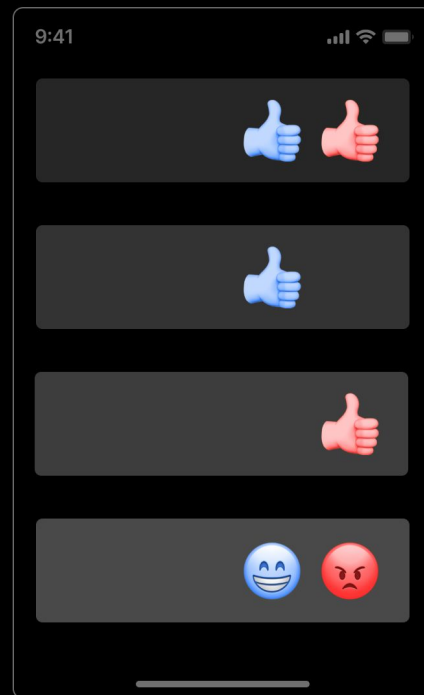


Context

ENGAGEMENT
based ranking
for [redacted]



BRIDGING
based ranking
for [redacted]

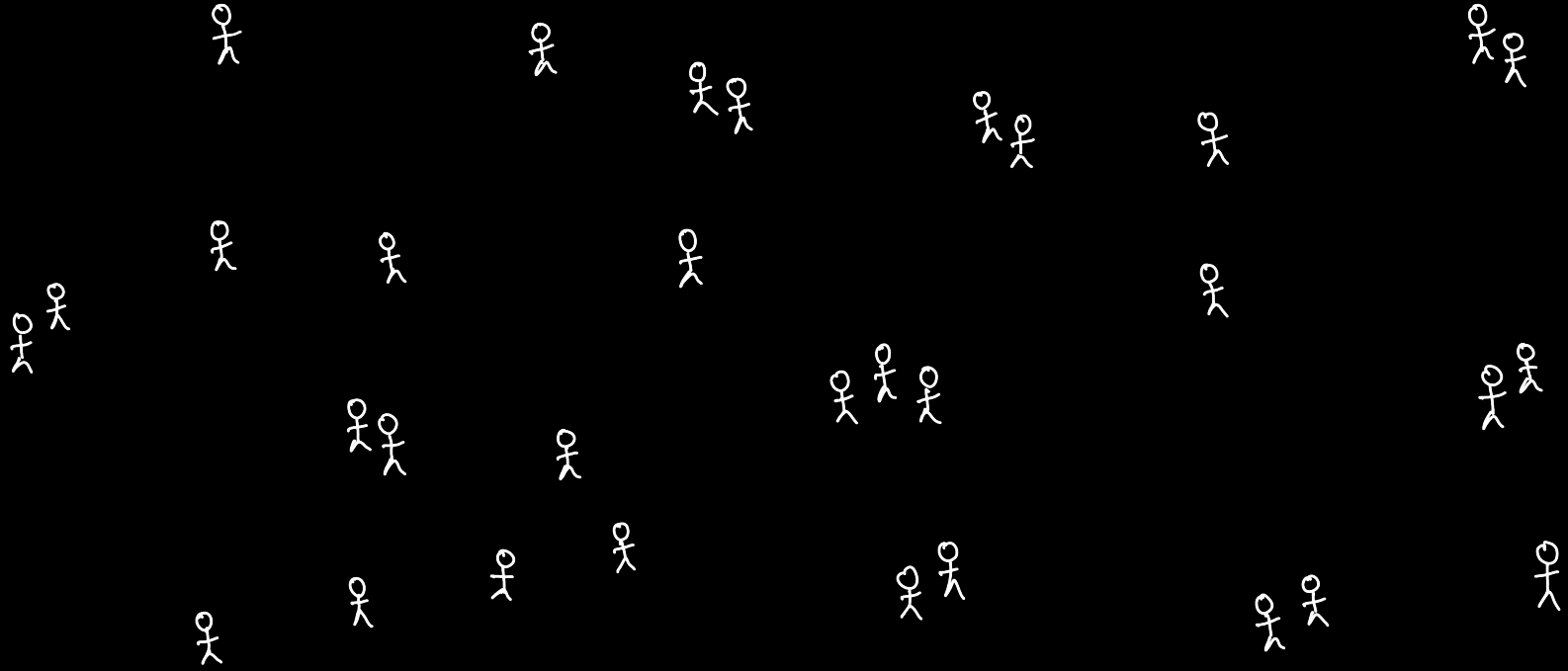


Disclaimer!

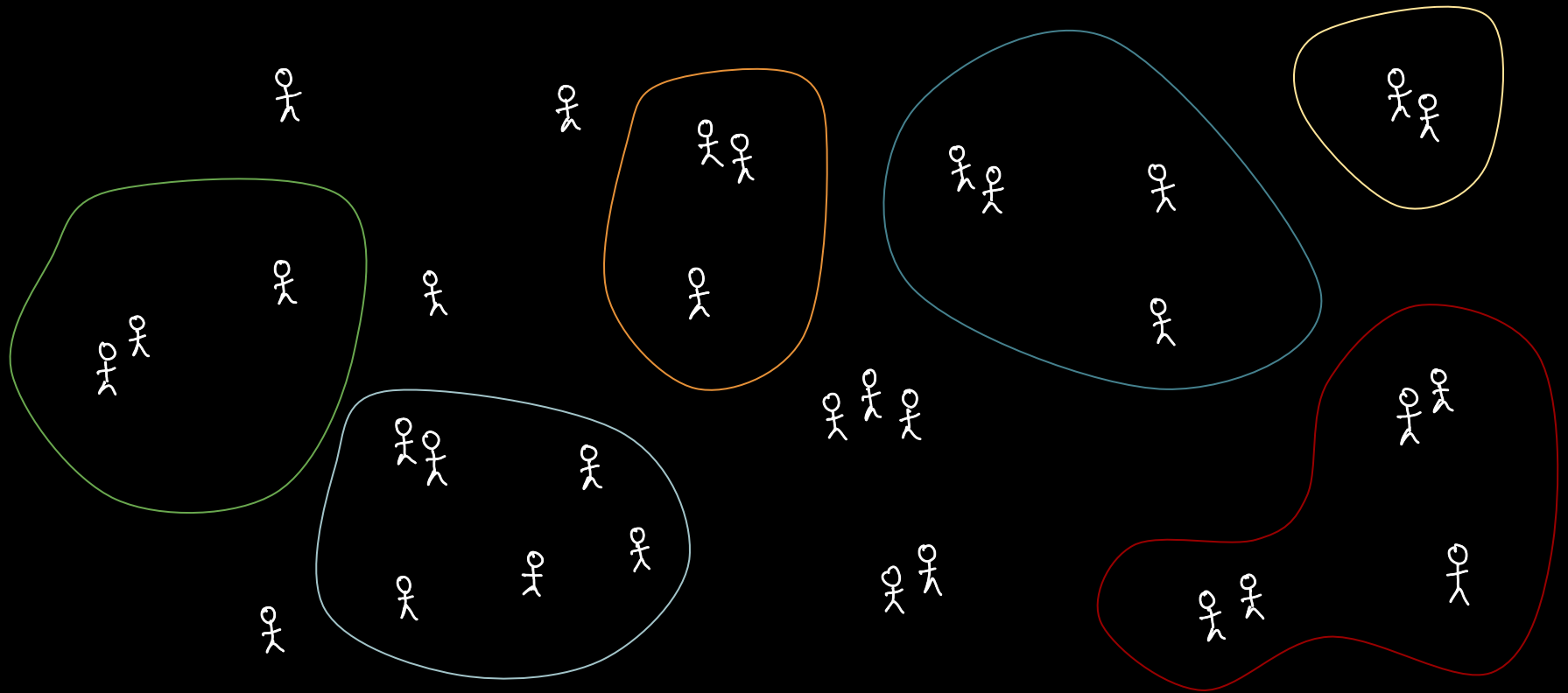
- I. Sortedness
- II. Trade-offs
- III. Goal (revisited)

- I. Sortedness
- II. Trade-offs
- III. Goal (revisited)

Societies are hypergraphs (intersecting groups).



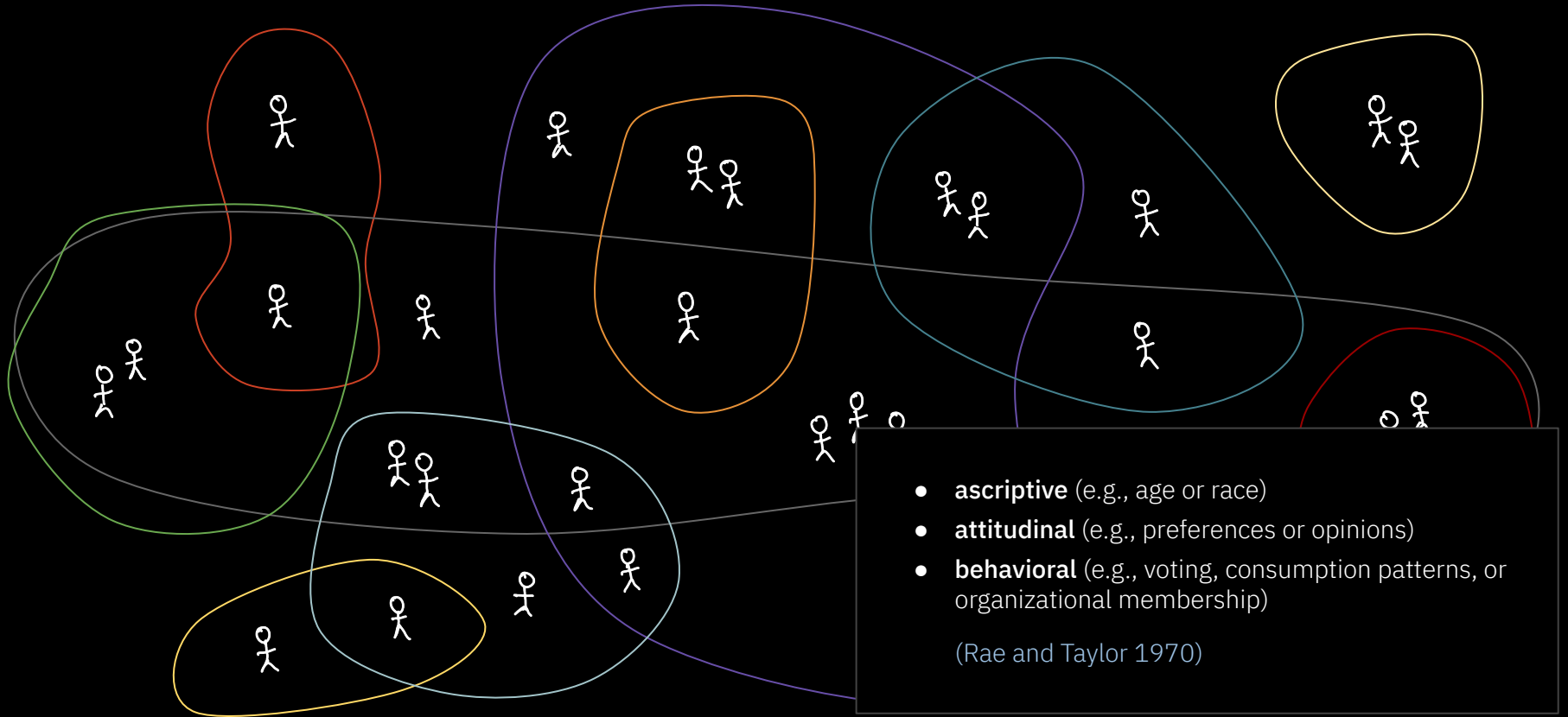
Societies are hypergraphs (intersecting groups).



Societies are hypergraphs (intersecting groups).



Societies are hypergraphs (intersecting groups).



Societies are hypergraphs (intersecting groups).



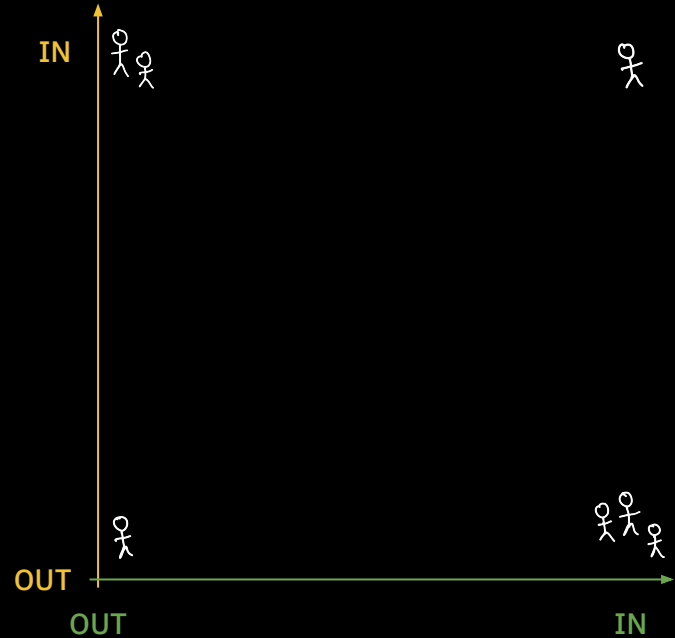
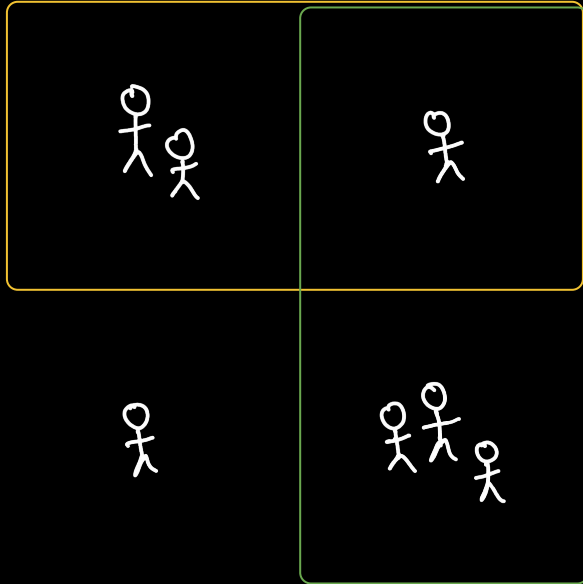
How to quantify the overall amount of crosscutting?



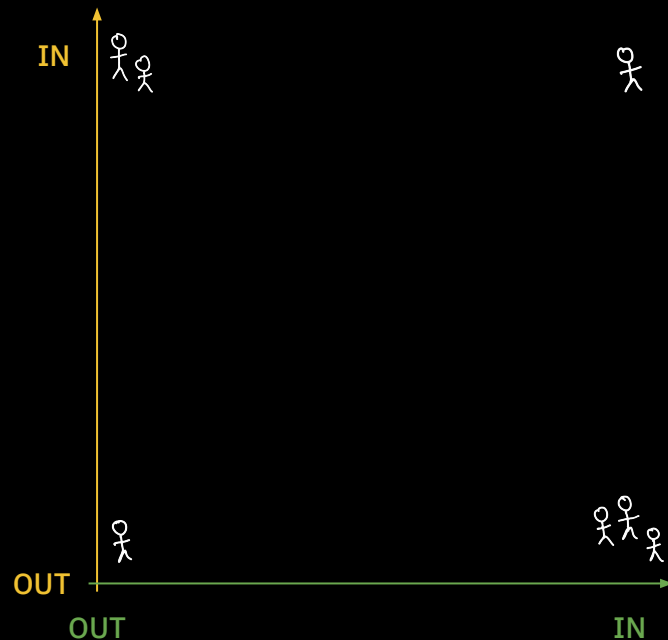
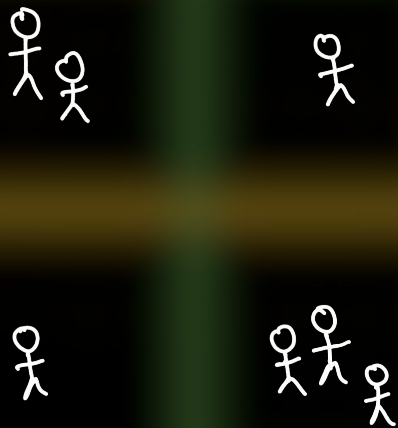
How to quantify the overall amount of crosscutting?



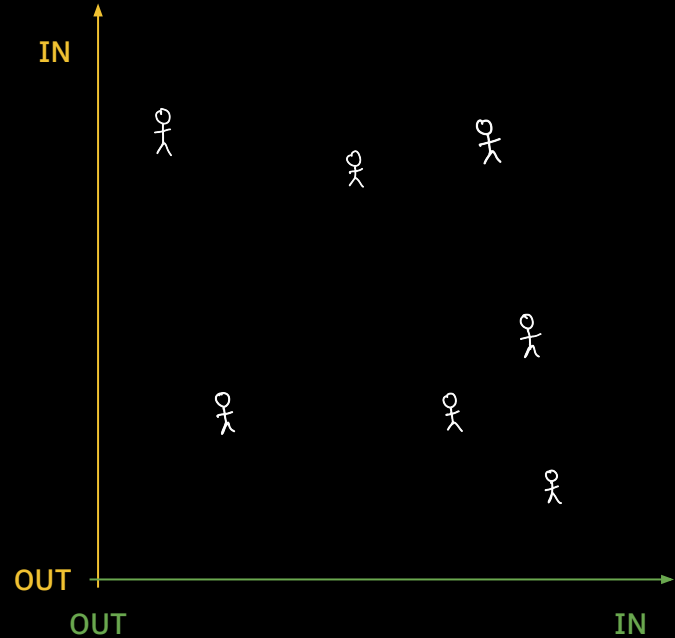
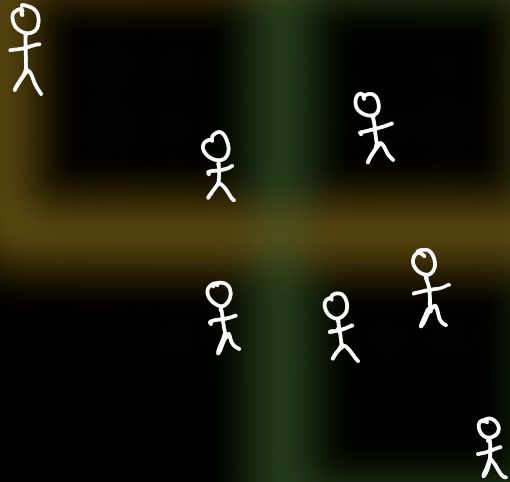
There is a correspondence between hypergraphs and distributions.



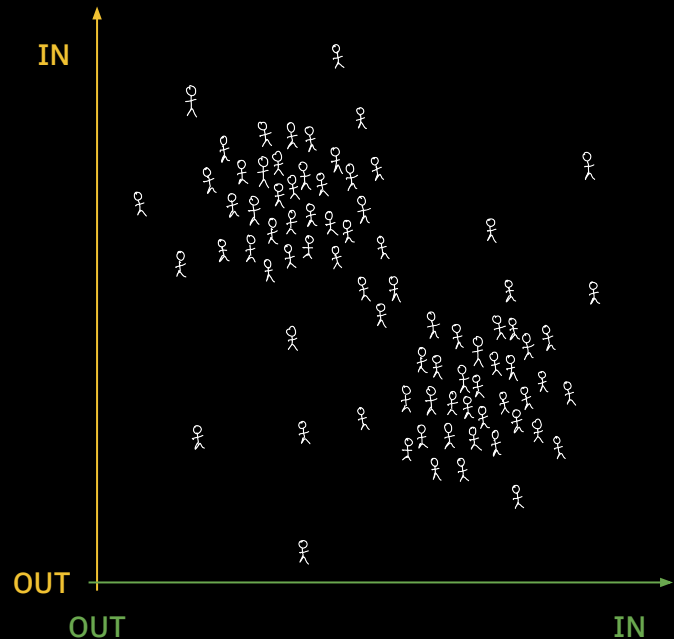
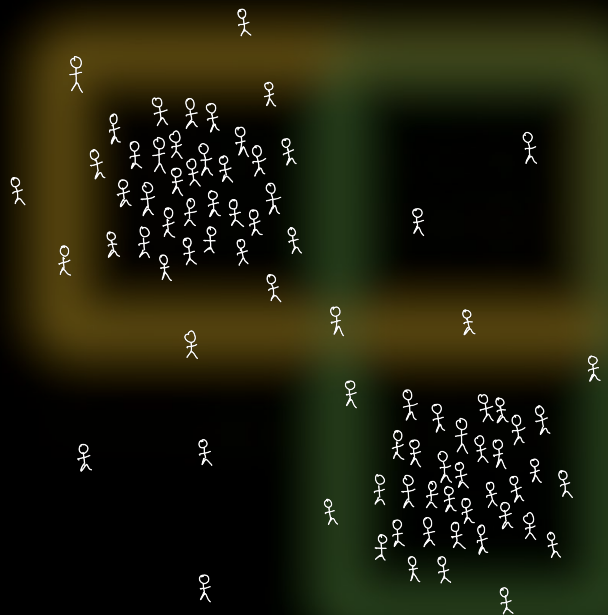
There is a correspondence between hypergraphs and distributions.



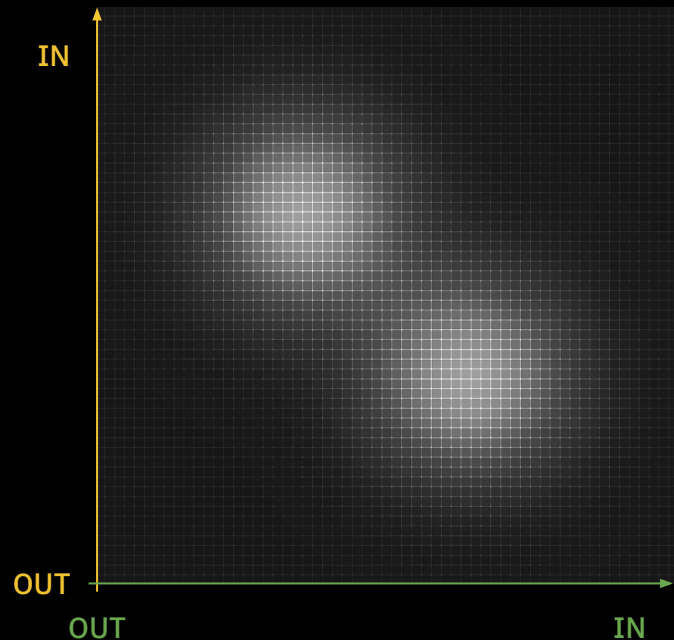
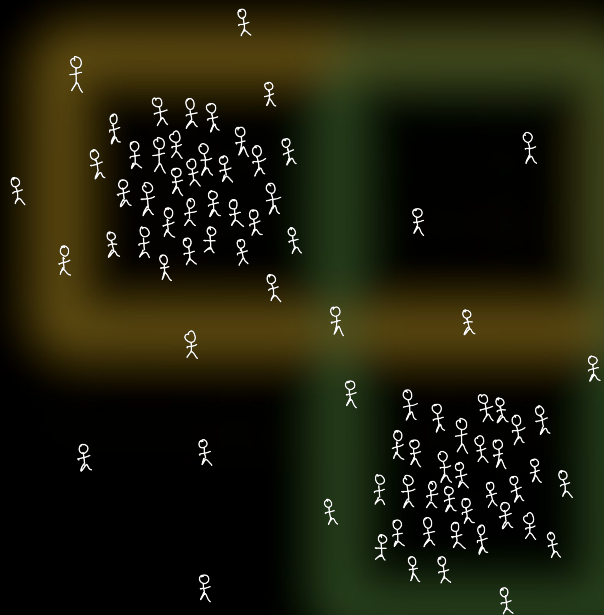
There is a correspondence between hypergraphs and distributions.



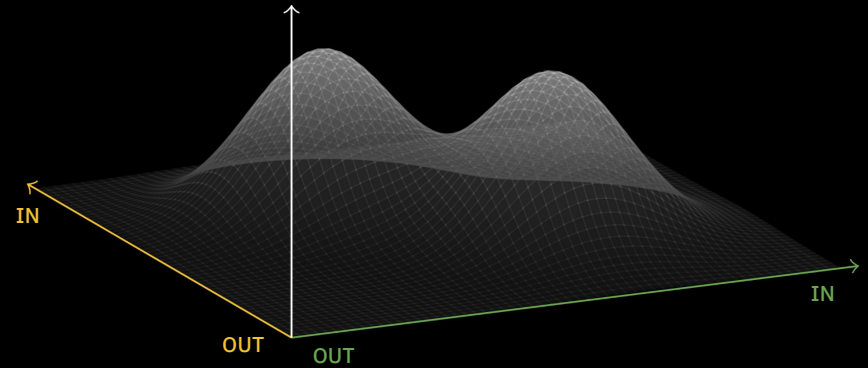
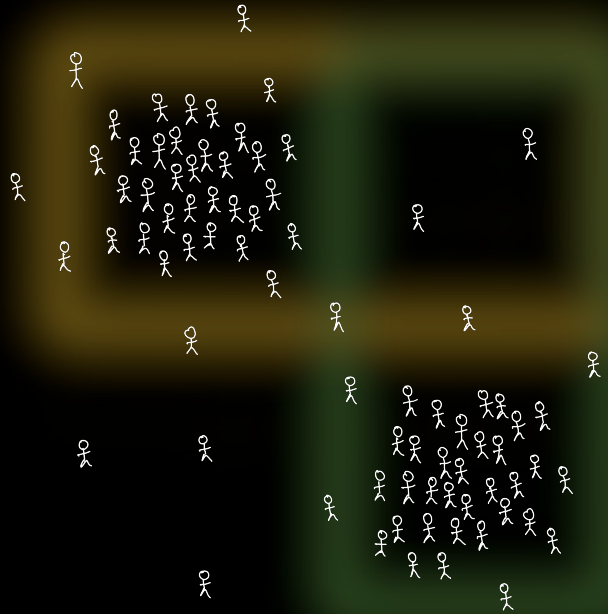
There is a correspondence between hypergraphs and distributions.



There is a correspondence between hypergraphs and distributions.

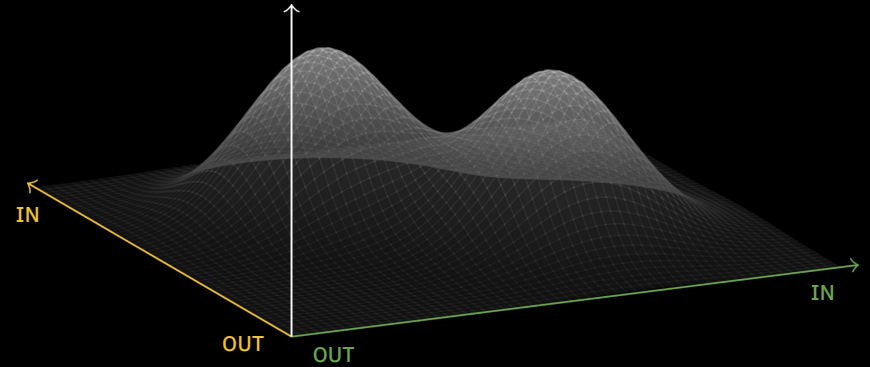
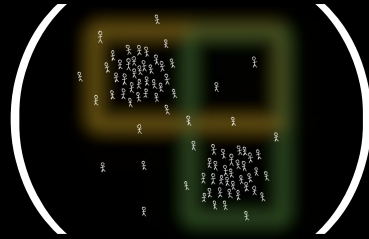


There is a correspondence between hypergraphs and distributions.



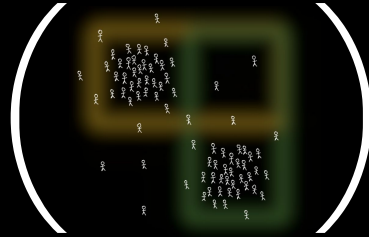
Crosscutting \propto Dependence

Crosscuttingness

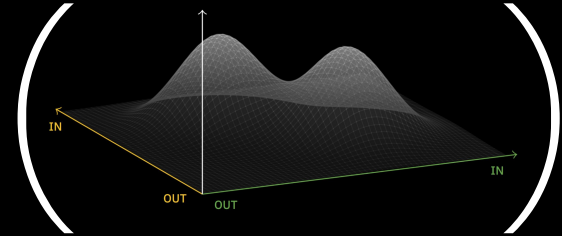


Crosscutting \propto Dependence

Crosscuttingness

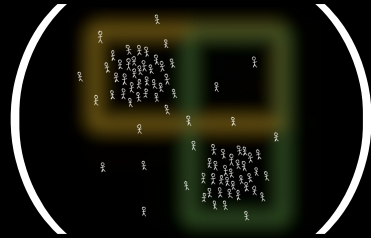


\propto Dependence

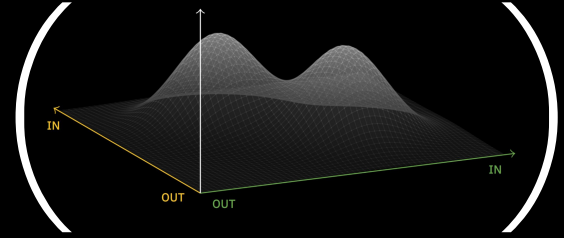


Crosscutting \propto Dependence = Sorting

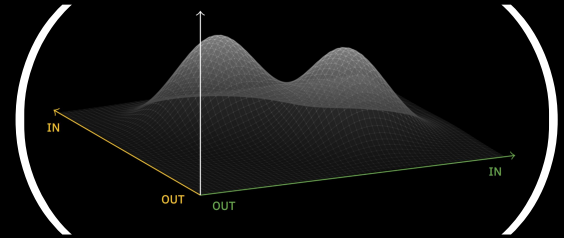
Crosscuttingness



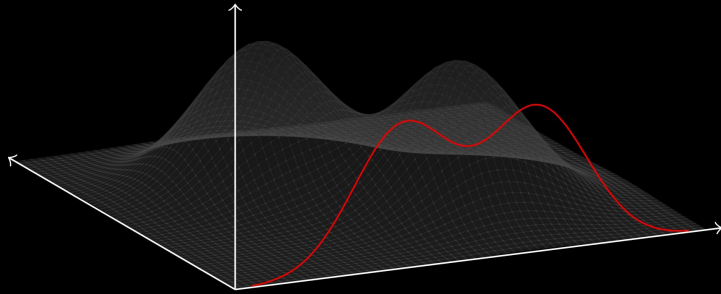
\propto Dependence



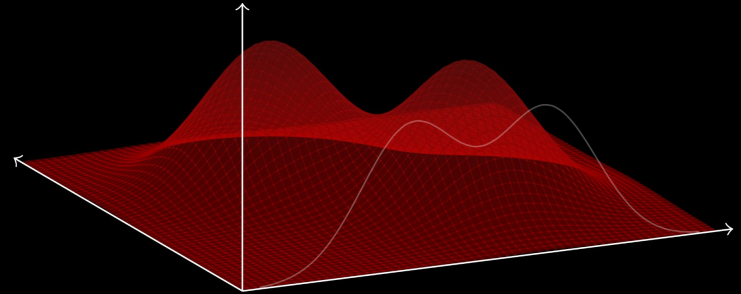
= Sorting



What is sorting?

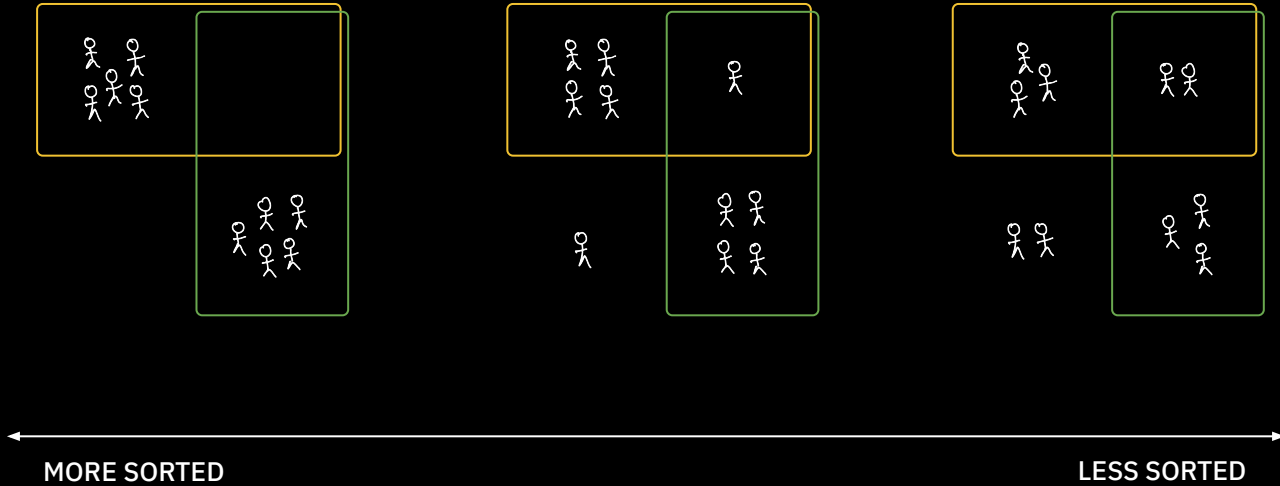


polarization

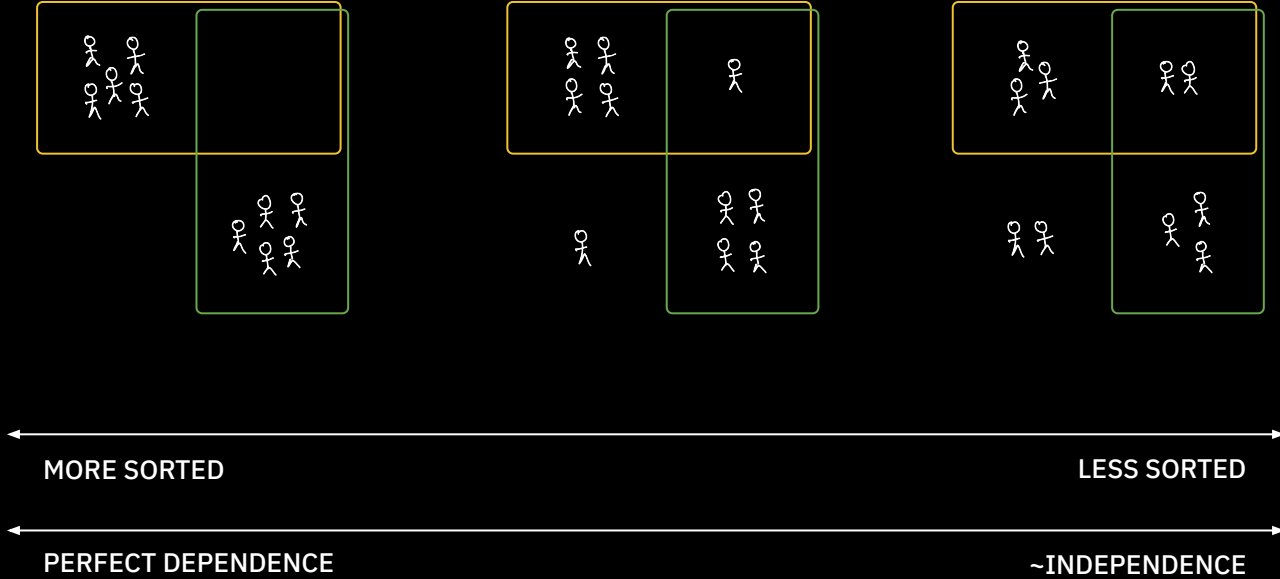


sorting

What is sorting?



What is sorting?

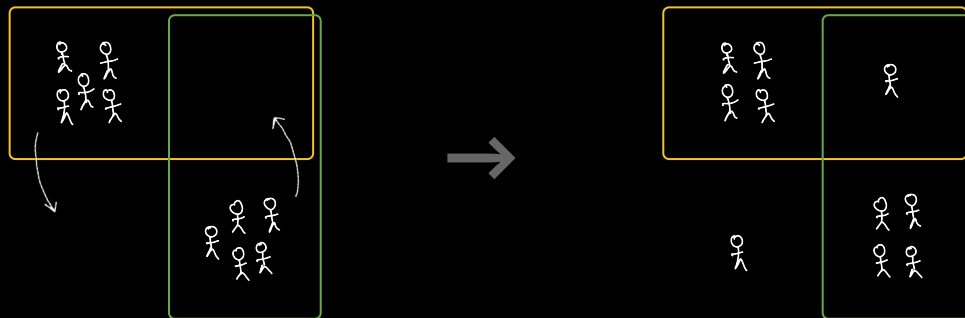


What is sorting?



Mechanism 1

Conversion



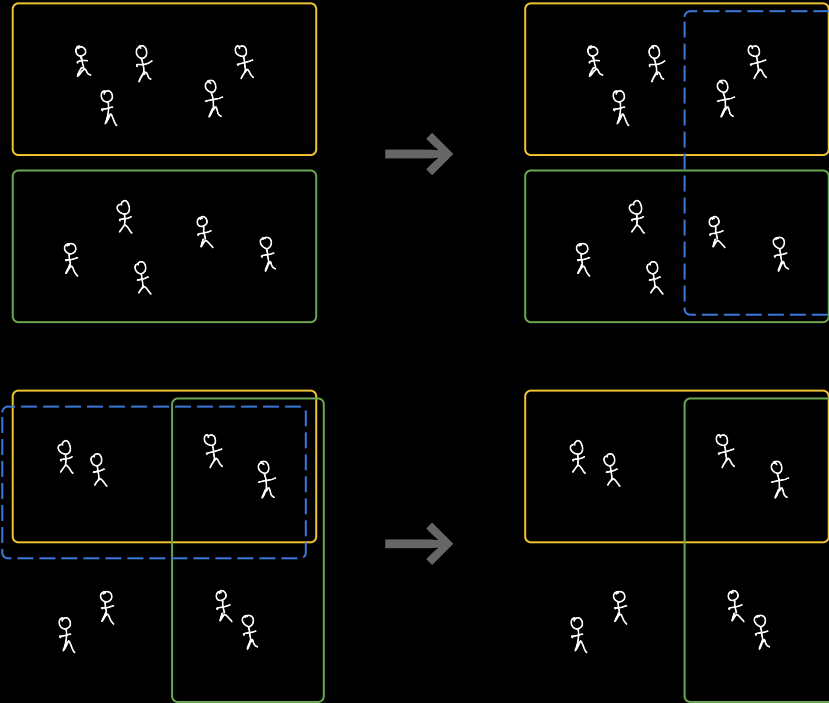
EXAMPLES

common ground
symmacy / common enemies
surprising validation

consilience
weak / long ties
“complicate the narratives”

Mechanism 2

Adding or removing groups/dimensions

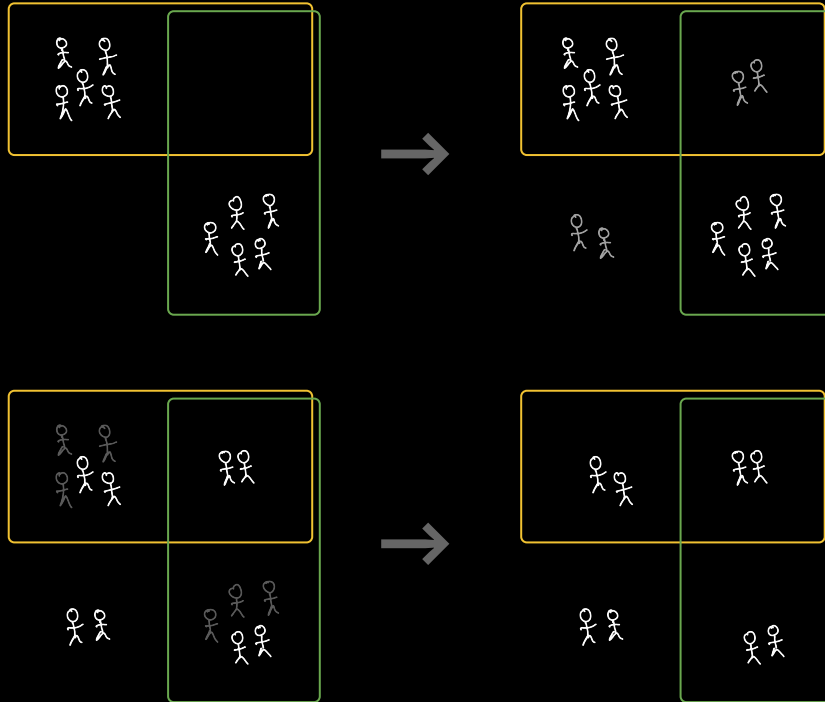


EXAMPLES

humanization
intergroup contact
depoliticization

Mechanism 3

Entry or exit of people



EXAMPLES

- migration
- generational change
- account creation / deletion

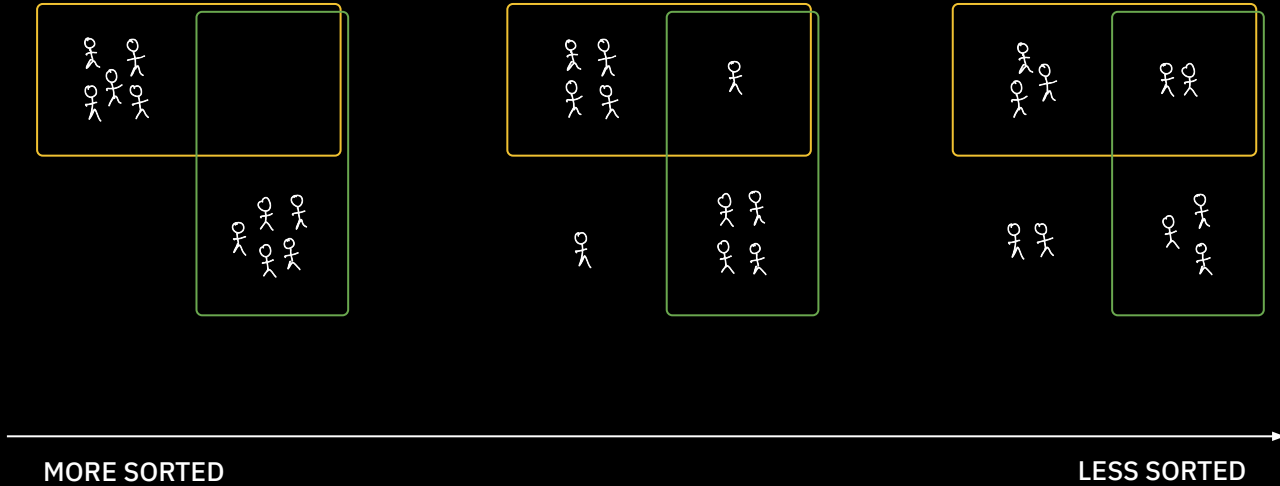
Most bridging algorithms use “diverse approval”.



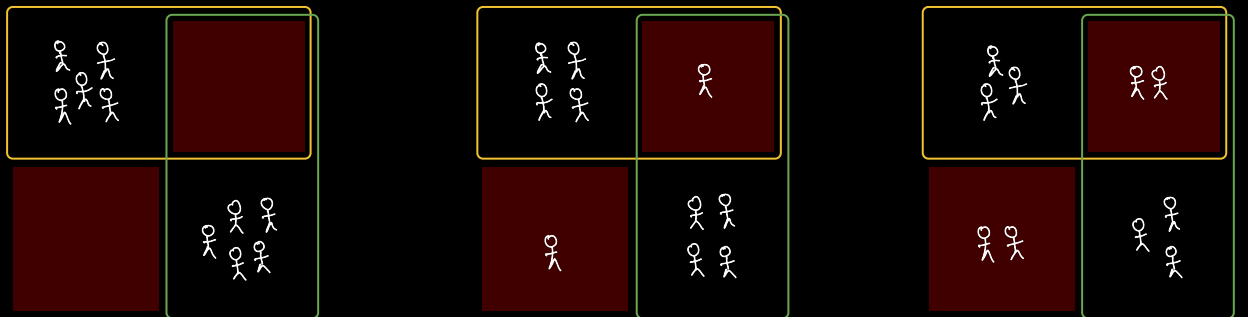
EXAMPLES



Diverse approval selects for unlikely combinations of attributes.



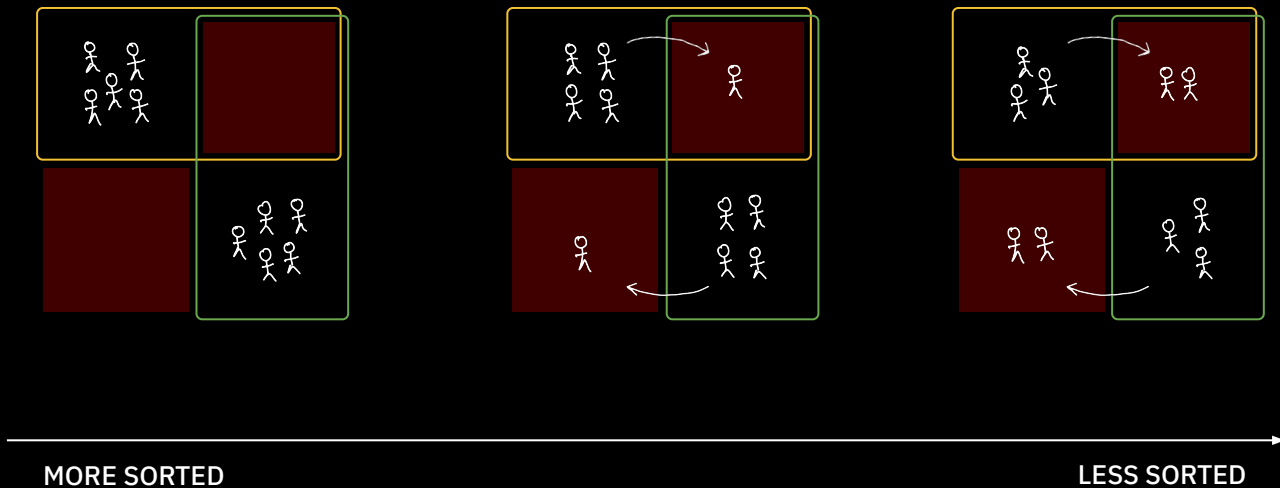
Diverse approval selects for unlikely combinations of attributes.



MORE SORTED

LESS SORTED

Diverse approval selects for unlikely combinations of attributes.



Diverse approval selects for unlikely combinations of attributes.

Cognitive “biases”

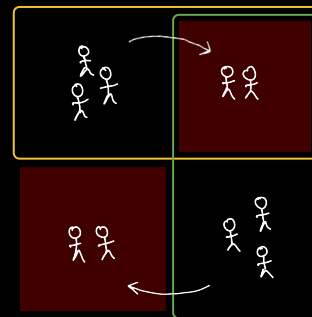
- mere-exposure effect
- illusory truth effect
- experience effects
- anchoring
- pressures to conform

Learning

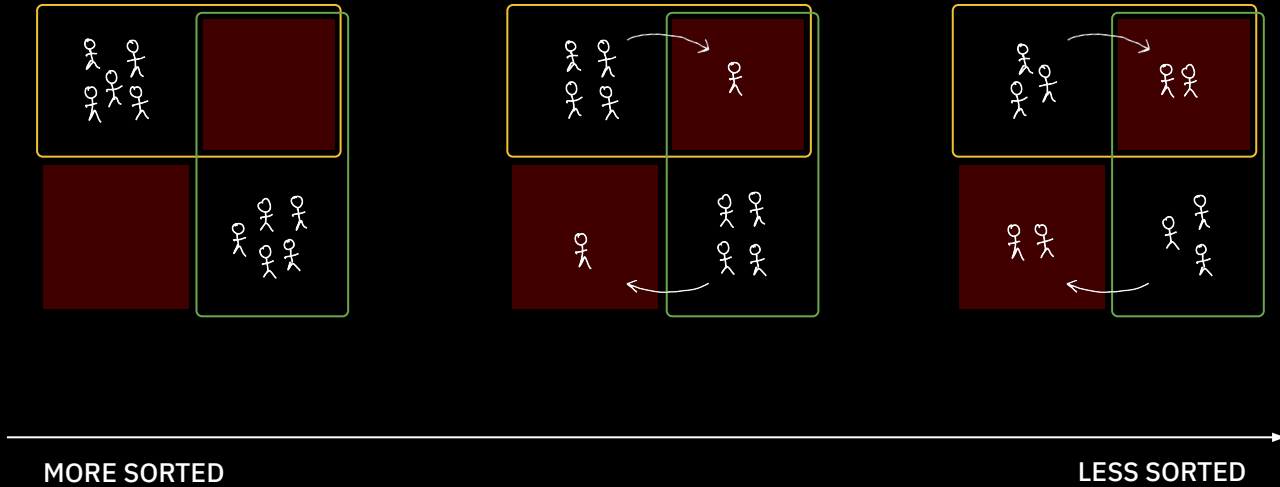
- learning new facts
- discovering new interests

Algorithmic incentives (Brady et al., 2023)

- observational learning
- reinforcement learning
- intentional strategic adaptation



Diverse approval selects for unlikely combinations of attributes.



- I. Sortedness
- II. Trade-offs
- III. Goal (revisited)

- I. Sortedness
- II. Trade-offs
- III. Goal (revisited)

Sorting is real.

Direct measurement:

(e.g., Abrams et al. 2015, Kaplan et al. 2022)

- ↑ correlations between issue positions
- “no evidence of polarization; the middle has not shrunk”
- ↑ geographic sorting

Explains empirical trends:

(Levendusky 2009; Abrams & Fiorina 2015; Fiorina 2017)

- ↑ difference in approval ratings
- ↓ split ticket voting
- campaigning to convince → campaigning to mobilize
- ↑ affective polarization

Sorting is bad.

Unsorting is good.

Sorting is real.

Direct measurement:

(e.g., Abrams et al. 2015, Kaplan et al. 2022)

- ↑ correlations between issue positions
- “no evidence of polarization; the middle has not shrunk”
- ↑ geographic sorting

Explains empirical trends:

(Levendusky 2009; Abrams & Fiorina 2015; Fiorina 2017)

- ↑ difference in approval ratings
- ↓ split ticket voting
- campaigning to convince → campaigning to mobilize
- ↑ affective polarization

Sorting is bad.

- ↑ pigeonholing / stereotyping
- ↓ representation of nuanced, crosscutting positions (Fiorina 2016)
- ↑ risk of civil war (Selway 2011; Gubler and Selway 2012; Siroky and Hechter 2016)

Unsorting is good.

Sorting is real.

Direct measurement:

(e.g., Abrams et al. 2015, Kaplan et al. 2022)

- ↑ correlations between issue positions
- “no evidence of polarization; the middle has not shrunk”
- ↑ geographic sorting

Explains empirical trends:

(Levendusky 2009; Abrams & Fiorina 2015; Fiorina 2017)

- ↑ difference in approval ratings
- ↓ split ticket voting
- campaigning to convince → campaigning to mobilize
- ↑ affective polarization

Sorting is bad.

- ↑ pigeonholing / stereotyping
- ↓ representation of nuanced, crosscutting positions (Fiorina 2016)
- ↑ risk of civil war (Selway 2011; Gubler and Selway 2012; Siroky and Hechter 2016)

Unsorting is good.

- ↑ “surprising validators” (Glaeser and Sunstein 2014)
- ↑ extent to which any majority will have something in common with a minority
- ↑ weak/long ties → ↑ economic outcomes (Jahani et al., 2023)

Sortedness by other names:

safeguards against factionalism

James Madison, *The Federalist No. 10* (1787)

“intersection of social circles”

Georg Simmel, *Soziologie* (1908)

“overlapping memberships”

Robert Dahl, *A Preface to Democratic Theory* (1956)

“overlapping consensus”

John Rawls, *Political Liberalism* (1993)

“diverse diversities”

Amartya Sen, *Identity and Violence* (2006)

“connected society” / “polypolitanism”

Danielle Allen, *Justice by Means of Democracy* (2023)

Unsorting is bad.

- too neutral a goal (e.g., QAnon)
- absurd if taken to the extreme
- may be useful for social capital, trust, and belonging
- makes engaging in politics more complicated
- may be manipulative

Diverse approval (may) have side-effects.

Unsorting is bad.

- too neutral a goal (e.g., QAnon)
- absurd if taken to the extreme
- may be useful for social capital, trust, and belonging
- makes engaging in politics more complicated
- may be manipulative

Diverse approval (may) have side-effects.

- may disadvantage minority or non-centrist perspectives
- may privilege those who are good at articulation
- may incentivize bothsidesism (false balance)
- may surface ideas with diverse but incompatible support (e.g., removing §230)

- I. Sortedness
- II. Trade-offs
- III. Goal (revisited)

- I. Sortedness
- II. Trade-offs
- III. Goal (revisited)

So what should the goal be?



Margin

Build a recommender that unsorts because it would be better than the status quo, so is justified (for now). Also, exists in a broader context so likely not able to overcome all the other tendencies towards divisiveness in society.

So what should the goal be?



Margin

Build a recommender that unsorts because it would be better than the status quo, so is justified (for now). Also, exists in a broader context so likely not able to overcome all the other tendencies towards divisiveness in society.



Mirror

Build a recommender that gives people an accurate understanding of the degree of sorting that is present in a population. (“Show people the bridging that already exists.”)

So what should the goal be?



Margin

Build a recommender that unsorts because it would be better than the status quo, so is justified (for now). Also, exists in a broader context so likely not able to overcome all the other tendencies towards divisiveness in society.



Mirror


Build a recommender that gives people an accurate understanding of the degree of sorting that is present in a population. (“Show people the bridging that already exists.”)





Anti-Manipulation


Build a recommender that has no systematic tendency towards sorting or unsorting.

So what should the goal be?

-  **Margin**


Build a recommender that unsorts because it would be better than the status quo, so is justified (for now). Also, exists in a broader context so likely not able to overcome all the other tendencies towards divisiveness in society.
-  **Mirror**


Build a recommender that gives people an accurate understanding of the degree of sorting that is present in a population. (“Show people the bridging that already exists.”)
-  **Anti-Manipulation**


Build a recommender that has no systematic tendency towards sorting or unsorting.
-  **Moderate Manipulation**


Build a recommender that unsorts, except in certain protected ways (or only in certain greenlit ways).


So what should the goal be?

-  **Margin**

Build a recommender that unsorts because it would be better than the status quo, so is justified (for now). Also, exists in a broader context so likely not able to overcome all the other tendencies towards divisiveness in society.
-  **Mirror**

Build a recommender that gives people an accurate understanding of the degree of sorting that is present in a population. (“Show people the bridging that already exists.”)
-  **Anti-Manipulation**

Build a recommender that has no systematic tendency towards sorting or unsorting.
-  **Moderate Manipulation**

Build a recommender that unsorts, except in certain protected ways (or only in certain greenlit ways).
-  **Mandate**

Build a recommender that unsorts, but only if it has a mandate to do so (e.g., via democratic decision or individual consent).

So what should the goal be?



Margin

Build a recommender that unsorts because it would be better than the status quo, so is justified (for now). Also, exists in a broader context so likely not able to overcome all the other tendencies towards divisiveness in society.



Mirror

Build a recommender that gives people an accurate understanding of the degree of sorting that is present in a population. (“Show people the bridging that already exists.”)



Anti-Manipulation

Build a recommender that has no systematic tendency towards sorting or unsorting.



Moderate Manipulation

Build a recommender that unsorts, except in certain protected ways (or only in certain greenlit ways).



Mandate

Build a recommender that unsorts, but only if it has a mandate to do so (e.g., via democratic decision or individual consent).

→ bridging.systems
→ lukethorburn.com



Aviv Ovadya
Berkman Klein



Maria Polukarov
KCL



Carmine Ventre
KCL



Jonathan Stray
CHAI



Smitha Milli
Cornell Tech



Helena Puig Larrauri
BuildUp

Thanks to the many collaborators & mentors who have contributed to or informed this work →



Seth Lazar
ANU



Emillie de Keulenaar
UN

This work was supported by UK Research and Innovation [grant number EP/S023356/1], in the UKRI Centre for Doctoral Training in Safe and Trusted Artificial Intelligence (safeandtrustedai.org), King's College London.