

O'Connor and Weatherall 2018: Polarization and Trust

Kevin Dorst

23.223, Rationality

I. Beliefs and Networks

O&W: argue that polarization (in politics and elsewhere) not necessarily caused by irrationality.

Starting point: scientists—and people generally!—figure out what to believe through two mechanisms:

- 1) The evidence *they* gather.
- 2) The evidence *others* gather.

⇒ Social networks are important to epistemology.

How? Hard to study. Build computational models.

II. Explore-exploit tradeoffs

Many repeated decisions involve a tradeoff between *gaining more information* about the problem (exploring), and *using your information* about the problem (exploiting).

Restaurant? Major? Dating?

Multi-armed bandits. Finite case. Myopic maximization vs. maximizing expected long-run reward. Getting stuck on wrong arm?
Infinite case.

Quickly gets complex!

Focus on simple two-armed case: Two options, A vs. B , with varying reward probabilities.

Say $ch(A = 1)$ known to be 0.5, while $ch(B = 1)$ is known to be either $0.5 + \epsilon$ or $0.5 - \epsilon$. Suppose B better.

III. Bala and Goyal model

Pieces:

- Two-armed bandit
- Bayesians in a social network.
- Choose action, pull some number of times, report results to neighbors; update beliefs; repeat.
- MEU even more intractable. Instead, myopic maximization.

People choose action, share evidence, repeat.

In this model, they always (eventually) *converge in opinions*.

Usually they converge on the truth (B better). But not always!

Why? Causes of ulcers; Palmer study.

Misleading evidence can throw everyone off!

Zollman effect: There is a tradeoff between *speed* and *reliability* of convergence.

More communication → faster;
Less communication → more reliable.
Why? The network itself exhibits an explore/exploit tradeoff.

IV. Modeling trust

Problem: in B&G model, people's beliefs always converge (unless disconnected network).

Not true in real social networks!

Polarization: persistent, large disagreements in beliefs between two or more factions. Eg Lyme wars

This is *one* kind of "polarization".

What drives polarization in Lyme wars?

- Not differences in values.
- Not "siloining"—all sides know the (basic) evidence on both sides.
- O&W: *differential trust*.

Introduce trust into B&G model: how much do you believe the evidence people share with you?

Jeffrey conditionalization on *E*. Which way (and how much) shift depends on trust.

Q: Is this a good model of trust?

What does trust depend on? O&C: "How similar their opinions are to yours: if people radically disagree with you, don't trust the evidence they share."

Result: polarization.

Conclusion: since differential trust is reasonable, polarization is to be expected from reasonable people.