



## Class 1

Sociological Theories and Agentbased Computational Models of Diversity and Polarization in Society Andreas Flache

Class lecture @ Latin American School and Workshop on Data Analysis and Mathematical Modeling of Social Science. SoFiA - SocioFisica Argentina November, 7-11, 2016 Buenos Aires, Argentina

## **Social influence**

#### The tendency to alter one's opinions, attitudes, beliefs, customs, or other cultural traits to more closely resemble those of influential others

"... Beliefs, attitudes, and behavior cover a wide range indeed, there are still more *things over which interpersonal influence extends*, such as language, art, technical standards, and social norms. The most generic term for the things over which people influence each other is culture."

"Culture is taken to be what social influence influences." Axelrod, Robert. 1997.

The Dissemination of Culture. Journal of Conflict Resolution.

# The universality of social influence?

- > Conformity experiments (e.g. Asch)
- > Small group research (e.g. Sheriff)
- Persuasion studies (e.g. Myers)
- > Innovation diffusion (e.g. Rogers)
- > Mass media research (e.g. Katz & Lazarsfeld)



## Axelrod's puzzle and Abelson's puzzle

## **Axelrod (1997):**

"If people tend to become more alike in their beliefs, attitudes, and behavior when they interact, why do not all such differences eventually disappear?"

## Abelson (1964)

"... What on earth one must assume in order to generate the *bimodal outcome* of community cleavage studies?"

Abelson RP. 1964. Mathematical models of the distribution of attitudes under controversy.

## **One form of diversity:**

## Local convergence and global diversity

### **Pluriformity in opinion distributions**

#### **Example: Political and social views in the US**

Glaeser & Ward (2007) analyzed data from PEW 1987–2003 Values Survey ( $\approx$  2500 respondents) and concluded

" America is a country with remarkable geographic diversity in its habits and beliefs. People in different states have wildly different views about religion, homosexuality, AIDS, and military policy, as well as wildly different consumption patterns...The extent and permanence of cultural divisions across space is one of America's most remarkable features."

Quoted from Glaeser, Edward L. and Bryce A. Ward. 2006. "Myths and Realities of American Political Geography", *Journal of Economic Perspectives* 20(2), pp. 119-144.

## Local convergence and global diversity: Kissing in France (yet another example)



Kissing business acquaintances X, XXX or XXXX?

## *Source: Website of The Economist, Oct 24th*

"...A colleague tells of a friend who found himself on the border between a three- and a two-kiss stronghold. In his words, the number of times you were expected to touch cheeks literally depended on which way you turned when leaving the house in the morning."

#### Source map: http://www.radicalcartography.net/

#### Another form of diversity: (bi-)polarization (DiMaggio et al. 1996, Evans 2003, Fischer et al 2009)

 increasing variance of opinion distribution in some issues in some historical periods (e.g. Sexual morality; attitudes
Trump and Clinton are neck and neck in Florida. The issue of immigration is a crucial — and divisive — issue in the state. No surprise, exit polls indicate that the candidates' voters have radically different views on the impact that immigrants have on the country today: 50 percent of Trump voters believe immigrants hurt the country, while a whopping 83 percent of Clinton voters say the opposite, that immigrants help.



## Traditional explanations: Individual differences

- > Age, region, education, race...
  - ⇒"different cultural tastes are concentrated within different sociodemographic segments" (Mark 2003)
- > Explanations: identity, status, networks...

### Remaining questions

- > How do similar people "coordinate" on a similar culture?
- > How is it possible that differentiation arises even within groups of similar people?

## Traditional explanations (2): Differentiation without sociodemographic differences

## Division of labor (functional differentiation, individualization)

- > Due to population growth, technological change
- Generates individual differences and structural differentiation (network patterns)

## Remaining question

Division of labor also implies social interaction and thus social influence

⇒ Why does influence not eliminate other differences?

# Computational models of social diversity

- > Models proposed by Carley, Axelrod, Mark, Latane...
- > Multiple agents
  - States: cultural "attributes", opinions
  - Relations: likelihood of interaction, strength of influence
- Homophily
  - the higher the similarity, the more likely the interaction (relational dynamic).
- > Influence:

if there is interaction, the interactants become more similar (state dynamic).

> Interaction & influence is restricted to local neighbors

## **Diversity and polarization as puzzle?**

**Classical models of social influence in networks** (e.g. French, Abelson, Harary, Lehrer & Wagner,...) *Social Influence*: move towards weighted average opinion of others to whom one is connected, weighted by strength of connection



## **Classical models: Pluriformity collapses into consensus**

- Incremental updating of continuous bounded opinion
- Actors opinion moves towards weighted average of neighbors' opinion

(Abelson 1964; Berger 1981; DeGroot 1974; French 1956; Harary 1959; Lehrer 1975)



In connected networks, opinions will always converge to perfect consensus

Graph taken from: Mäs M, Flache A (2013). PLoS ONE 8(11): e74516.

## Axelrod's explanation of "local convergence and global diversity"

Interaction of **two** fundamental social mechanisms

"birds who flock together sing the same song"  $\Rightarrow$  social influence "birds of a feather flock together"  $\Rightarrow$  homophily

Plus:

- Discrete (nominal) cultural states
- Local interaction (cellular grid)



Axelrod, R. 1997. The Dissemination of Culture. *Journal of Conflict Resolution* 41:203-226.

## Further aligning Axelrod with sociological theory: from interpersonal to social influence In Axelrod-type models influence is *interpersonal*

 $\Rightarrow$  single "deviant" is highly influential

> but social influence is group phenomenon.



Flache & Macy (2011, JCR) integrate *social* influence

Agent adopts trait supported by largest number of *influential* neighbors

- Who is influential?
- $\Rightarrow$  Local interaction + homophily + noise

## Social influence more in detail ...

- 1. Agent is chosen at random
- **Every** neighbor is influential with probability *p* = *overlap* (= homophily as in original Axelrod)
- 3. Agent adopts trait supported by largest number of *influential* neighbors (social influence!)

keep current trait if there is a tie.

## (More) robust global diversity in Axelrod framework



Model assumptions:

- local interaction
- > social influence
- > homophily
- > noise
- > large radius
- > large population

Flache, A. & M.W. Macy 2011. Local Convergence and Global Diversity: From Interpersonal to Social Influence. *Journal of Conflict Resolution*.

## ("Real") social influence



means of 50 replications)

## Answers to Axelrod's puzzle?

- Local interaction
- > Discrete (nominal) opinion space

Given there is noise in the world:

- Culture-network co-evolution (network homophily)
- > "Layered" social influence
- "Real" social influence

Conditions for robust diversity:

Limited noise, population size, complexity 'cultural space' ...

## More challenges: metric features and blending

- Axelrod etc assume nominal opinion space
  - Either you agree or you don't: direction and degree of influence on an issue can not be expressed
- > *Metric scaling* may often be more adequate
  - "What should be the age at first marriage"
  - Traditional opinion formation models (French, Abelson...)
- > Metric features allow *"blending"*: gradual compromising



## Importing metric scaling and blending in the Axelrod framework (≅ Flache, Macy, 2011, JMS)

Like earlier continuous models (French etc):

Social Influence:

Move towards average opinion of influential neighbors j

$$q_{iF,t+1} = q_{iF,t} + \frac{1}{C} \sum_{j \in neigh} W_{ijt} (q_{jF,t} - q_{iF,t})$$

Unlike these earlier models:

Homophily:

The more overlap *i*-*j*, the more influence does *j* have

$$w_{ij,t+1} = 1 - \frac{\sum_{f=1}^{F} \left| q_{ift} - q_{jft} \right|}{\text{MaxDist}}, \quad 0 \le w \le 1$$

## **Back to monoculture**

Why metric features promote monoculture

- As long as distance not maximal, two agents remain connected.
- > Maximal distance is very unlikely from random start
- $\Rightarrow$  Back to French etc earlier result:

monoculture in connected network



F=5, Q=10.000, N=32x32, rad 6, No noise

Class 1 Models of Diversity and Polarization – Flache – SoFiA 2016

### Adding homophily to social influence: "Bounded confidence"

- > Homophily (Lazarsfeld & Merton ...)
  - interact more likely with more similar others, be more open to influence from more similar others
- > "Bounded confidence" models
  - Hegselmann, Krause, Weissbuch, Deffuant, ...
  - If disagreement is too large, agents do not influence each other any more ("confidence threshold").
- Bounded confidence assumes that influence goes to zero if differences are too large.

## Bringing Axelrod's solution back in: Interaction thresholds

## "Bounded confidence" models

- > Weissbuch, Deffuant, Hegselmann, Krause...
- > If disagreement is too large, agents do not influence each other any more ("confidence threshold").

**Resurrects Axelrod's "homophily" assumption that influence goes to zero if differences are too large** 



## **Interaction threshold: Implementation**

- Interaction threshold
  - if opinion similarity is too small (threshold τ), probability of interaction drops to zero
  - Otherwise it corresponds to average overlap, as before.

$$w_{ij,t+1} = \begin{cases} \sum_{j=1}^{F} |o_{ijt} - o_{jjt}| \\ 1 - \frac{\sum_{j=1}^{F} |o_{ijt} - o_{jjt}|}{\text{MaxDist}}, & \text{if this is } \geq \text{threshold} \\ 0, & \text{otherwise} \end{cases}$$



## Effects of changing the threshold (1)



Class 1 Models of Diversity and Polarization – Flache – SoFiA 2016



## Effects of changing the threshold (2)



Medium threshold

(τ=.85)



Source: Hegselmann & Krause 2002, JASSS, p. 10

Class 1 Models of Diversity and Polarization – Flache – SoFiA 2016



## Effects of changing the threshold (3)



Source: Hegselmann & Krause 2002, JASSS, p. 10

- The BC-model can explain stable diversity and clustering
- However, small perturbation of opinions (white noise) is a problem.

(Mäs, Flache & Helbing 2010, PLoS Comp Biol)



## Noise destroys diversity in BC model

## One possible explanation of robust pluriformity: negative influence

- Positive influence: change opinion to move closer to position of influential others.
- > Negative influence: change opinion to move away from position of influential others.
- > Two versions of negative influence:
  - The more similar others are, the more likely one moves away.

## • The more dissimilar others are, the more likely one moves away.

## Noise destroys diversity in BC model

## One possible explanation of robust pluriformity: negative influence

- Positive influence: change opinion to move closer to position of influential others.
- > Negative influence: change opinion to move away from position of influential others.
- > Two versions of negative influence:
  - The more similar others are, the more likely one moves away.
    - The more dissimilar others are, the more likely one moves away.

#### The more similar others are, the more likely one moves away: "striving for uniqueness"

- "there are two opposing forces, the one centripetal, the other centrifugal", Durkheim 1897 (2003a: 258-259)
  - <u>Integrating forces</u> motivate individuals to conform with contacts (social influence)
  - <u>Disintegrating forces</u> motivate individuals to deviate from social norms (individualization)
    - > When individuals perceive that too many others hold a similar opinion, they adjust their opinions.
    - Supported by psychological research on striving for uniqueness (Imhoff et al. 2009; Maslach et al. 1985; Snyder et al. 1980)



#### The model

- > Computational ABM of opinion dynamics
- > At each simulation event, an agent is picked and opinion (*oi*) is updated

$$\Delta o_{i} = \frac{\sum_{j=1}^{N} \left( w_{i,j} \left( o_{j} - o_{i} \right) \right)}{\sum_{j=1; i \neq j}^{N} w_{i,j}} + x$$

Integrating forces (Social influence)

Disintegrating forces (white noise)



> Influence weights  $(w_{ij})$  are positive and decrease as opinion distance increases (homophily, similarity-attraction)

$$w_{i,j} = e^{-\frac{dist_{ij}}{A}}$$

The higher **A** the less are agents influenced by individuals with different opinions.

 Individual changes (x) are modeled as white noise. Standard Deviation of noise depends on similarity to all other agents

$$x \sim N(0, s \sum_{j=1}^{N} e^{-dist_{i,j}})$$

**s** is used to vary the strength of the disintegrating forces

Class 1 Models of Diversity and Polarization – Flache – SoFiA 2016

9-11-2016



#### Typical run with little "disintegrating force": Consensus



- > A=2, s=.4
- > Integrating forces prevail above disintegrating forces
- > No perfect consensus but very small opinion variance
- Individual perturbations add up to collective opinion shifts (random walk)



Typical run with very strong disintegrating force: Anomie



- → *A*=2, *s*=6
- > Disintegrating forces prevail above integrating forces
- > Black line shows trajectory of one agent



#### **Clustering with intermediate disintegrating force**



More details:

Mäs, Flache & Helbing 2010, PLoS Comp Biol Mäs, Flache & Kitts 2014, Perspectives on Culture and Agent-based Simulations

## Clustering as equilibrium state



B: high initial opinion variance, uniform distribution (sd=0.3)



C: maximal opinion polarization at the outset

Intermediate strength disintegrating force

#### Source:

Mäs, Flache & Kitts 2014, Perspectives on Culture and Agent-based Simulations

Mäs, Flache & Helbing 2010, PLoS Comp Biol.



tion – Flache – SoFiA 2016

#### Conditions for clustering



Cluster=

set of agents in adjacent positions such that each set member was separated from the adjacent set of members by a maximum of 5 scale points

> 9-11-2016



## So far...

- Pluriformity of opinions (clustering) is not inconsistent with negative influence of the "striving for uniqueness" type, if
- the relative strength of positive social influence and "striving for uniqueness" are in balance.
- $\Rightarrow$  Relatively small window of the parameter space
- ⇒Pluriformity is then an absorbing state of the system, independent of initial opinion distribution.

#### The other type of negative influence:

#### The more dissimilar others are ... xenophobia and negative influence

Agents distance themselves from disliked others, dissimilar others are disliked

- social balance theory, cognitive dissonance theory, optimal distinctiveness theory, and
- empirical evidence for "negative referents", "profiling"

Various models include xenophobia and negative influence

(Macy, Kitts, Flache, Benard 2003, see also Mark 2003, Jager & Amblard 2004, Baldassari & Bearman 2007, Flache & Mäs 2008, Fent, Groeber & Schweitzer 2007, ...)

Xenophobia

• if opinion difference too large, relations become negative Negative influence

• If relations are negative, agents increase opinion distance

## Modelling the interplay of positive and negative influence

Extending earlier social influence models (e.g. Flache & Macy 2011, JMS) *Positive influence (assimilation) and negative influence (differentiation):* 

• local neighbours "pull" or "push" opinion o depending on weight  $w_{ij}$ 

$$\Delta o_{i,t} = \frac{1}{\text{scalingfactor}} \sum_{j \neq i} w_{ij,t} (o_{j,t} - o_{i,t}) \qquad 0 \le o \le +1$$

*Homophily and xenophobia*: change of relational weight *w* 

• average distance > zero  $\Rightarrow$  positive weight, else *negative* 



## Interplay of positive and negative influence A typical result: initial pluriformity turns into polarization



## **Pluriformity is also possible ...**

- > Initially unimodal random. s.d. 0.25
- > N=10000, 250000 iterations
- Synchronous updating

Pluriformity can persist when for every subgroup attraction to "friends" and rejection from "foes" exactly balance



Class 1 Models of Diversity and Polarization – Flache – SoFiA 2016

200 150 100

opinion

1 44

frequency

opinion

# Pluriformity is a highly fragile state in between consensus and polarization

Pluriform opinion distributions are highly fragile



Initial distribution is stable pluriform

⇒ Change only one agent's opinion very slightly

⇒ Collapse into polarization





## Whether consensus or polarization occurs depends (amongst other things) on demographic similarities or faultlines (segregation) between groups



## A condition for polarization?

- Lau and Murnighan (1998): "group splits" may be caused by "strong demographic faultlines"
- Strong faultline: high correlation of different demographic attributes

	Attribute	Member A	Member B	Member C	Member D		
	Gender	male	male	female	female		
Group 1	Age	20	40	20	40		
	Skin color	black	white	white	black		
	Gender	male	male	female	female		
Group 2	Age	20	20	40	40		
	Skin color	white	white	black	black		



#### Modelling effects of "diversity": Integrating demographic features (Flache & Mäs 2008 CMOT, SimPat)

Similarity i-j depends on both demographic and opinion (dis)similarity:

$$w_{ij,t+1} = 1 - \frac{2\left(\sum_{d=1}^{D} \left|s_{id} - s_{jd}\right| + \sum_{f=1}^{F} \left|o_{ift} - o_{jft}\right|\right)}{\text{MaxDist}}, \quad -1 \le w_{ij,t+1} \le 1$$

D fixed demographic features s,

F changing opinion features o

## A strategy to avoid polarization IF there is negative influence: "timing of contacts"

To avoid polarization under strong faultline, initially form demographically homogenous subgroups

- ⇒ within separate subgroups there is mainly demographic similarity: positive influence prevails
- ⇒ most agents are not extreme initially. Subgroups coordinate on similar "moderate local norms"
- ⇒ once that has happened, positive influence will also prevail in contact between subgroups. *Flache & Mäs, 2008, CMOT*

## **Modelling faultline strength**

#### Implementation of faultline strength

 $A_2$ 

.8

 $A_1$ 

 $A_2$ 

 $A_3$ 

.8

.8

 $A_2$ 

.6

 $A_1$ 

 $A_2$ 

 $A_3$ 

.6

.6

	f = 0.9			<i>f</i> = 0.8			f = 0.7				f = 0.6				f = 0.5				
i	attr.	attr.	attr.		attr.	attr.	attr.		attr.	attr.	attr.		attr.	attr.	attr.		attr.	attr.	attr.
1	1	1	1	Î I	1	1	1	1	1	1	1	[	1	1	1		1	1	1
2			.1						1										
2		-1	-1			-1	-1		-1	-1	-1			-1	-1		-	-1	-1
3			1			1			1										1
4	-1	-1	-1		-1	-1	-1		-1	-1	-1		-1	-1	-1		-1	-1	-1
2		1	1			1	1		1	1	1			1	1		1	1	1
6	-1	-1	-1		-1	-1	-1		-1	-1	-1		-1	-1	-1		-1	-1	1
7	1	1	1		1	1	1		1	1	1		1	1	-1		1	1	-1
8	-1	-1	-1		-1	-1	-1		-1	-1	1		-1	-1	1		-1	-1	1
9	1	1	1		1	1	-1		1	1	-1		1	1	-1		1	1	-1
10	-1	-1	1		-1	-1	1		-1	-1	1		-1	-1	1		-1	-1	1
11	-1	-1	-1		-1	-1	-1		-1	-1	-1		-1	-1	-1		-1	1	-1
12	1	1	1		1	1	1		1	1	1		1	1	1		1	-1	1
13	-1	-1	-1		-1	-1	-1		-1	-1	-1		-1	1	-1		-1	1	-1
14	1	1	1		1	1	1		1	1	1		1	-1	1		1	-1	1
15	-1	-1	-1		-1	-1	-1		-1	1	-1		-1	1	-1		-1	1	-1
16	1	1	1		1	1	1		1	-1	1		1	-1	1		1	-1	-1
17	-1	-1	-1		-1	1	-1		-1	1	-1		-1	1	1		-1	1	1
18	1	1	1		1	-1	1		1	-1	-1		1	-1	-1		1	-1	-1
19	-1	1	-1		-1	1	1		-1	1	1		-1	1	1		-1	1	1
20	1	-1	-1		1	-1	-1		1	-1	-1		1	-1	-1		1	-1	-1
Σ	0	0	0		0	0	0	1	0	0	0	L	0	0	0	1	0	0	0

 $A_2$ 

.4

 $A_1$ 

 $A_2$ 

 $A_3$ 

.4

.4

 $A_2$ 

.2

 $A_1$ 

 $A_2$ 

 $A_3$ 

.2

.2

 $A_2$ 

.0

 $A_1$ 

 $A_2$ 

 $A_3$ 

.0

.0

16 1 50

## **Effects of faultline strength**

- > D=3 fixed features and K=4 "opinions"
- > Opinions initially randomly distributed maximal faultline does not make polarization trivial

$$w_{ij,t+1} = 1 - \frac{2\left(\sum_{d=1}^{D} \left|s_{id} - s_{jd}\right| + \sum_{f=1}^{F} \left|o_{ift} - o_{jft}\right|\right)}{\text{MaxDist}}, \quad -1 \le w_{ij,t+1} \le 1$$

#### **Effects of faultline strength**



Fig. 4. Effect of faultline strength on outcome measures, averages over 500 replications per conditions, outcomes measured after 1000 iterations per replication N = 20, D = 3, K = 4.

Class 1 Models of Diversity and Polarization – Flache – SoFiA 2016

## **Computational experiment: timing**

- > How would it effect opinion dynamics if we form initially subgroups?
- > In particular, what would happen if we initially keep highly dissimilar actors apart and "merge" subgroups only later?

## **Modelling timing of contacts**



## Results: right timing avoids negative faultline effects

500 replications per condition, outcomes measured after 100 iterations per replication N=20, D=3, F=4

Source: Flache, A, M. Mäs. 2008. How to get the timing right. Computational and Mathematical Organization Theory 14.1:23-51.



## What goes really on at the microlevel?

### **Controlled lab experiments**

RESEARCH ARTICLE

Discrepancy and Disliking Do Not Induce Negative Opinion Shifts

2016

Károly Takács<sup>1</sup>\*, Andreas Flache<sup>2</sup>, Michael Mäs<sup>2</sup>

OPEN O ACCESS Freely available online

PLOS ONE

PLOS ONE

#### Differentiation without Distancing. Explaining Bi-Polarization of Opinions without Negative Influence

Michael Mäs<sup>1</sup>\*, Andreas Flache<sup>2</sup> 2013

1 Chair of Sociology, in particular Modeling and Simulation, ETH Zurich, Zurich, Switzerland, 2 Department of Sociology/ICS, University of Groningen, Groningen, The Netherlands

## "Discrepancy and Disliking Do Not Induce Negative Opinion Shifts"



Experiments show: Influence mainly positive

No more negative influence if large disagreement

Takács, Flache & Mäs Plos One 11(6): e0157948. doi:10.1371/journal.pone.0157948



## Can we explain polarization and faultline effects without negative influence?

Yes – we can!

With a model based on persuasive argument theory (Mäs, Takács, Flache & Jehn, 2012, *Org Science; Mäs & Flache, 2013, PloS One*)

#### But this model has some very different implications

## Intergroup polarization without negative influence

A model based on persuasive argument theory (Mäs, Flache, Takács & Jehn, 2013, Organization Science; Mäs & Flache, 2013 PlosOne)

- > Opinion is constituted by **arguments** arg\_vector ++---- ⇒ opinion = -0.33
- > **Homophily**: the more similar, the more likely interaction
- > **Influence**: if *i* interacts with *j*, then *i* adopts argument from *j*.



#### $\Rightarrow$ interaction with similar others increases polarization

### Persuasive argument theory: Opinion polarization with maximal faultline



Dynamics of opinion and interaction network

#### with maximal faultline

Further assumptions:

- strong homophily
- demographically biased opinions

Source: Mäs, Flache, Takács & Jehn, 2013, Organization Science

### Persuasive argument theory: Consensus with slightly weaker faultline



Now we added one (!) "criss-crossing" actor (all other things equal)

- ⇒ Sooner or later arguments communicated between opposing subgroups
- $\Rightarrow$  System moves into consensus eventually

Source: Mäs, Flache, Takács & Jehn, 2013, Organization Science

## Topic of class 2: **Ø PL** Testing theories of cooperation nduce and polarization Differ Polari in the Lab. Michael Mi 1 Chair of Sociole

Netherlands