A Marketing Game:

A Reinforcement Learning Approach to Optimizing Preference on a Social Network

Matthew G. Reyes

O'Reilly Applied AI
April 18, 2018

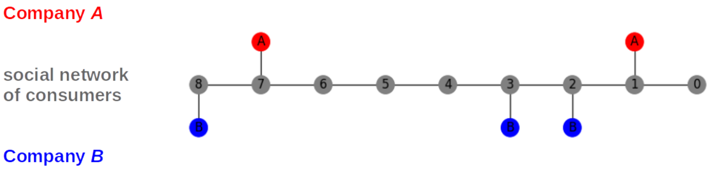
Motivation and Contribution

- consumers choose between two alternatives, A and B
  - Pepsi vs. Coke
  - Donald vs. Hillary
- preference modeled w/ socially contingent random utility
  - probabilistic utility maximization [McFadden '74]
  - utility depends on preferences of social connections [Blume '93]
- **Contribution of this work:**
  - re-parametrize model to incorporate influence of marketer
  - provides an operational approach to influencing preference

A Marketing Game

- social network of consumers
- competition between marketers to influence preference between two alternatives: Product A and Product B

$$x_i = \begin{cases} 1 & \text{if consumer } i \text{ prefers } A \\ -1 & \text{if consumer } i \text{ prefers } B \end{cases}$$



Brief Outline

- Psychology of Choice (Preference)
- Inferring States of Mind (Preference) from Data
- Graphical Model using Inferred States

Psychology of Preference

Why Consider this Problem?

- important from an intellectual point of view: import to understand influences on our decision-making
- marketers seek to influence our preferences in favor of their product or political candidate
- a model for influencing social decision-making could potentially be used to detect such attempts by adversarial governments
- a *market* is a set of alternatives from which *consumers* choose

Emphasis of This Approach

- seek to understand the influences that consumers exert upon one another's decision-making
- such information can be useful in resource allocation
  - perhaps you cannot influence someone directly, but you can influence someone who already exerts influence over them

Models of Choice: Differences in Perceived Utility

- *law of comparative judgment* [Thurstone 1927]
  - preference based on perceived difference in quality
- *independence of irrelevant alternatives* (IIA) [Luce 1959]
  - relative selection of two alternatives not affected by a third
- *aspect elimination* [Tversky 1972]
  - sequential selection of features possessed by alternatives
  - introduced to address situations where IIA does not hold
- *prospect theory* [Kahneman and Tversky 1979]
  - perceived utility often based on risk avoidance
- *random utility* [McFadden 1974]
  - utility is maximized, but has a random component
  - random component subsumes utility based on status or risk
  - correlation of random components determines choice structure

Utility has an Unknown Random Component

- *random utility* [McFadden '74] states that utility assigned to an alternative includes random components

$$U = \begin{bmatrix} u_A + \epsilon_A \\ u_B + \epsilon_B \end{bmatrix}$$

  - u_A and u_B are *known* sources of utility
  - ϵ_A and ϵ_B are *unknown* sources of utility
  - with respect to a given market, choices will be influenced by factors external to market that the modeler does not know

Utility as Parametrization of Observed Choice Frequencies

- decompose utilities u_A and u_B according to information that can be collected, i.e.,

$$u_A = \sum_i \theta_i f_i$$

  where the f_i are factors thought to be important in influencing perceived value

- examples of f_i include *cost*, *current events*, *possible reward*

- fit parameters associated with observed data

Assumptions on Unknonwn Sources of Utility

- *random utility* [McFadden '74]: consumers maximize utility, probability of choosing Product A becomes

$$p(u_A + \epsilon_A > u_B + \epsilon_B) \;\; = \;\; p(\epsilon_A - \epsilon_B > u_B - u_A) \; .$$

- if the unknown utilities ϵ_A and ϵ_B are distributed as the *maxima* of sequences of i.i.d. variables, *and* unknown sources of utility are uncorrelated, then get *logit* choice model

$$p(A) \;\; = \;\; \frac{e^{u_A}}{e^{u_A} + e^{u_B}} \; ,$$

- different assumptions on ϵ_A and ϵ_B lead to different choice rules

Inherent Bias Towards Products

- α_i is *inherent bias* representing the difference in utility assigned to the two alternatives by consumer i

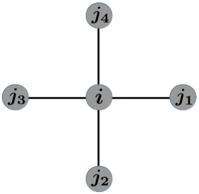
- probability of consumer i choosing alternative x_i is

$$p_i(x_i) \;=\; \frac{\exp\{\alpha_i x_i\}}{\exp\{\alpha_i x_i\} \;+\; \exp\{-\alpha_i x_i\}} \;=\; \frac{\exp\{\alpha_i x_i\}}{Z_i}$$

- also referred to as the *Luce* model [Luce '59]

Social Biases from Neighbors

- utility that consumer i assigns to alternatives A and B at time t is contingent upon choices $\mathbf{x}_{\partial i}^{(t)}$ of i's neighbors ∂i



- probability of consumer i choosing alternative x_i at time t given by Glauber dynamics

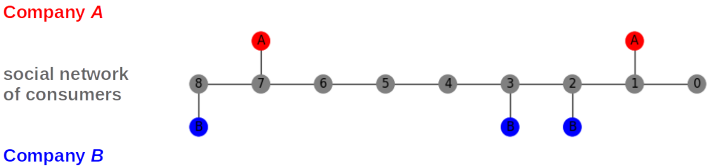
$$p(x_i|\mathbf{x}_{\partial i}^{(t)}) \quad = \quad \frac{\exp\left\{\sum_{j\in\partial i}\theta_{j\to i}x_i x_j^{(t)} + \alpha_i x_i\right\}}{Z_{i|\mathbf{x}_{\partial i}^{(t)}}}$$

where $\theta_{j\to i}$ is the *social bias* exerted upon i by j

Marketing Biases from Companies

- advertising by company influences the utility that consumers assign to alternatives

$$p(x_i|\mathbf{x}_{\partial i}^{(t)}) \quad = \quad \frac{\exp\left\{\sum_{j\in\partial i}\theta_{j\to i}x_i x_j^{(t)} + (\alpha_i + m_A^i - m_B^i)x_i\right\}}{Z_{i|\mathbf{x}_{\partial i}^{(t)}}}$$



Intuitive Interpretation of Our Model

- in *The Tipping Point*, Gladwell discussed factors responsible for the spread of ideas / preferences on a social network:
  - **salesmen** persuade others to purchase a product
  - **mavens** convince others with their expertise
  - **connectors** put people in touch with others
  - product **stickiness** keeps people coming back for more

$$p(x_i | \mathbf{x}_{\partial i}^{(t)}) = \frac{\exp\left\{ \sum_{j \in \partial i} \theta_{j \to i} x_i x_j^{(t)} + (\alpha_i + m_A^i - m_B^i) x_i \right\}}{Z_{i | \mathbf{x}_{\partial i}^{(t)}}}$$

Social Contagion: Spread of Preference

- others have considered socially-contingent decision-making in context of social contagion, spread of innovations, e.g.,
  - Kempe et al, 2005
  - Watts and Dodds 2007
  - Montanari and Saberi, 2010
- these works have considered best-response dynamics, i.e., a $\beta \to \infty$ scaling

$$p(x_i|\mathbf{x}_{\partial i}^{(t)}) = \frac{\exp\left\{\beta\left[\sum_{j\in\partial i}\theta_{j\to i}x_i x_j^{(t)} + \alpha_i x_i\right]\right\}}{Z_{i|\mathbf{x}_{\partial i}^{(t)}}}$$

- NOTE: no marketer!

Best-Response Good In Some Cases

- best-response amounts to selecting $\max\{u_A, u_B\}$

- corresponds to markets where "unknown" sources of utility are unimportant, i.e.,

$$p\left(\beta u_A + \epsilon_A > \beta u_B + \epsilon_B\right) \quad = \quad p\left(u_A - u_B > \frac{\epsilon_B - \epsilon_A}{\beta}\right) \ .$$

- makes sense when choices correspond to social / behavioral norms in which "fitting in" outweighs other considerations

Inferring Preference From Data

Random Utility Models are Data-Driven

- if we want to influence decision-making, must have a model that allows us to learn how individuals are making decisions
- random utility (exponential) models will fit parameters to observed factors so that resulting probability model predicts observed frequencies of choice
- any application will require experimentation with different parametrizations

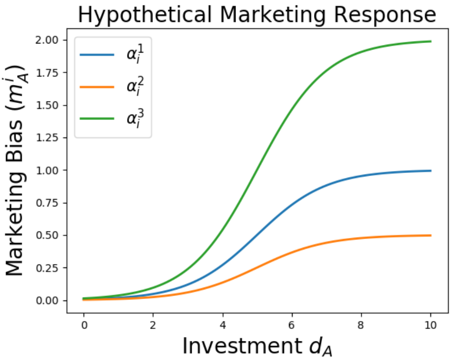
Marketer in Model Permits Reinforcement Learning

- **sensing:** *learn* direct and social biases $\{\theta_i\}$ and $\{\theta_{j\to i}\}$ with graphical model inference algorithms
- **reward:** seek to optimize *market share*
- **action:** *select marketing allocation* based on optimizing market share

$$p(x_i|\mathbf{x}_{\partial i}^{(t)}) = \frac{\exp\left\{\sum_{j\in\partial i}\theta_{j\to i}x_i x_j^{(t)} + (\alpha_i + m_A^i - m_B^i)x_i\right\}}{Z_{i|\mathbf{x}_{\partial i}^{(t)}}}$$

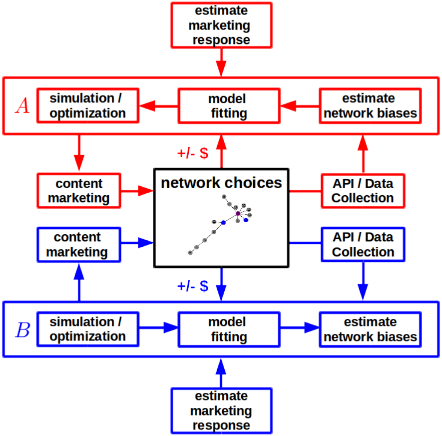
Marketing Strength as Function of Investment

- each consumer has a *marketing response* indicating their perception of value as a function of marketing intensity
- marketing response is with respect to type of marketing



High-Level Diagram

- learn influences from data; combine market research; simulate network model to select allocation

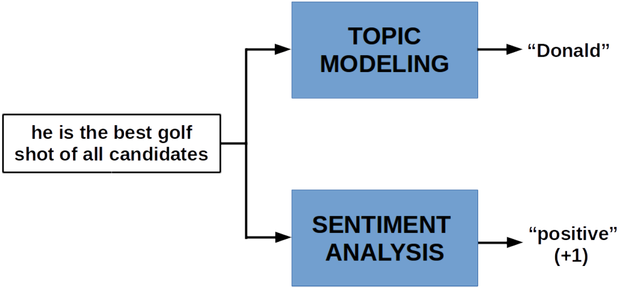


Deep Learning and Affective Computing

- consumer preferences determine data posted on social media
- consumer i will create post $y_i^{(t)}$ that is correlated with preference $x_i^{(t)}$
- deep learning, topic modeling, and sentiment analysis will infer semantic content of posts
- *affective computing* will infer preference state $x_i^{(t)}$ from semantic content of $y_i^{(t)}$
    - related to *theory of mind* psychology

Infer Preferences from Social Media Data

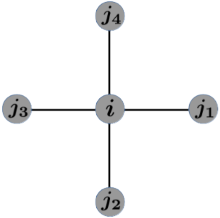
- apply machine learning algorithms to infer preferences of consumers from text / images shared on social media
  - deep learning, topic modeling to infer content
  - sentiment analysis, affective computing to infer attitude



Database of Preference Estimates

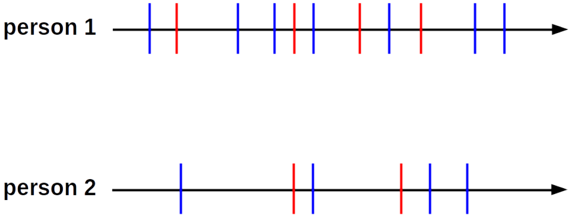
- applying machine learning to posted data yields *states* with respect to the choice problem under consideration, i.e., preference for Product A or Product B
- once we have estimated states, apply *graphical model* estimation algorithms to learn inherent and social biases, model expected behavior

t	i	j1	j2	j3	j4
0	1	1	-1	1	-1
1	1	1	1	-1	1
2	-1	-1	1	1	-1
...					



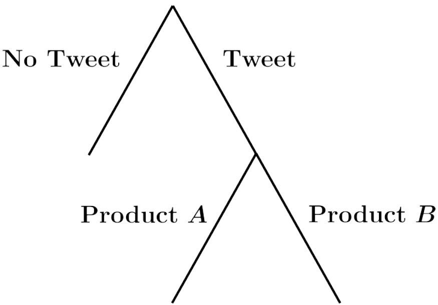
Users Who Tweet at Different Rates

- in paper, we assume that all users "update" their preference (post data) at the same rate



Nested Logit for Different Tweet Rates

- in paper, we assume that all users "update" their preference (post data) at the same rate



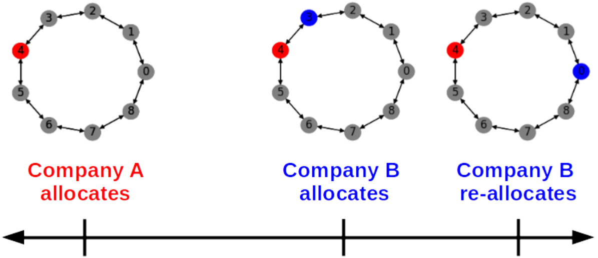
Graphical Model Problem

Properties of Social Networks

- small-world networks
- scale-free networks
- let's consider a cycle: allows us to simplify

Simplified Scenario

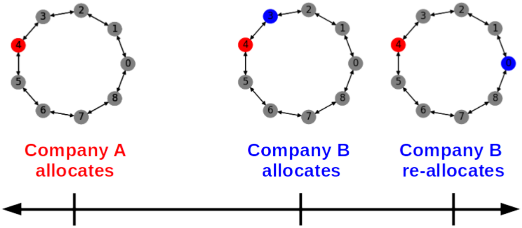
- each Company has one unit of (equal 'strength') marketing
- companies A and B take turns (re-)allocating



- specifically, we consider Company B's parameter estimation and allocation decision following Company A's allocation

Current Setting

- for all consumers i
  - $\alpha_i = 0$
  - $\theta_{i+1\to i} = 1$
  - $\theta_{i-1\to i} = .6$
- Company A allocates to consumer 4 with marketing strength $m_A^4 = 2$
- we will analyze steps in Company B's allocation selection



Asymmetric Glauber Dynamics

- if $\theta_{j\to i} = \theta_{i\to j} = \theta_{ij}$, Glauber dynamics converge to Gibbs equilibrium [Blume '93]

$$p(\mathbf{x};\theta) \quad = \quad \frac{1}{Z(\theta)} \exp\{\sum_{\{i,j\}} \theta_{ij} x_i x_j + \sum_{i \in V} \theta_i x_i\}$$

- Godreche showed that for a cycle with
  - $\theta_{i\to j} = \theta \pm \Delta$
  - $\theta_i = \theta'$

  Glauber dynamics converge to a stationary distribution that coincides with the symmetric Gibbs equilibrium

Reinforcement Learning Approach to Influencing Consumer Decision-Making [R, Computing Conference '19]

- expected total preference over sequence $\mathbf{x}^{(t_1)}, \ldots, \mathbf{x}^{(t_K)}$

$$r\left(\mathbf{x}^{(t_1)}, \ldots, \mathbf{x}^{(t_K)}\right) = \sum_{t=t_1}^{t_K} \sum_{i \in V} x_i^{(t)}$$

- Company A selects allocation M_A by simulating network dynamics based on estimated parameter $\hat{\boldsymbol{\theta}}^{(t_0)}$, candidate allocation M_A, and estimated preference configuration $\hat{\mathbf{x}}^{(t_0)}$

$$Q_{\hat{\mathbf{x}}^{(t_0)}}(\hat{\boldsymbol{\theta}}^{(t_0)}, M_A, \gamma, T) \triangleq \mathbf{E}\left[\sum_{\tau=0}^{T} \gamma^{\tau} r\left(\mathbf{X}^{(t_0+1+\tau)}\right) \middle| \hat{\mathbf{x}}^{(t_0)}\right]$$

Tracking Network Biases by Minimizing Conditional Description Length

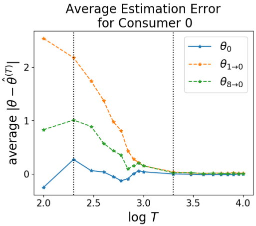
- estimate neighborhood parameters $\hat{\theta}_{\bar{i}} = \hat{\theta}_i \cup \{\theta_{j \to i}\}, j \in \partial i$ by minimizing *conditional* description length [R and Neuhoff]

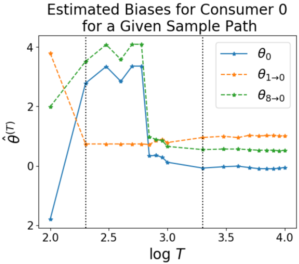
$$\bar{D}(x_i^{(t-T:t)} | \mathbf{x}_{\partial i}^{(t-T:t)}; \theta_{\bar{i}}) = -\sum_{\tau=1}^{T} \log p(x_i^{(t-\tau)} | \mathbf{x}_{\partial i}^{(t-\tau)}; \theta_{\bar{i}})$$

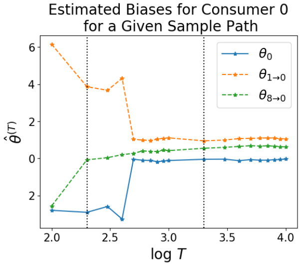
- mathematically equivalent to maximizing *pseudo* likelihood or performing logistic regression where the site preference is the response and the preferences of neighbors are the predictors; however, MCDL provides a more sound development

Transient vs. Stationary Phase

- equations work in the stationary phase
- most real-world problems will be transient

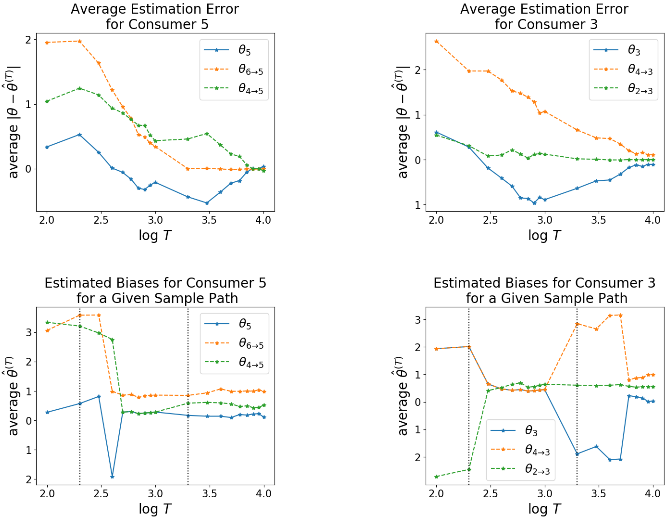






Tracking Direct and Social Biases

- recall that site 3 is influenced more by site 4 than is site 5

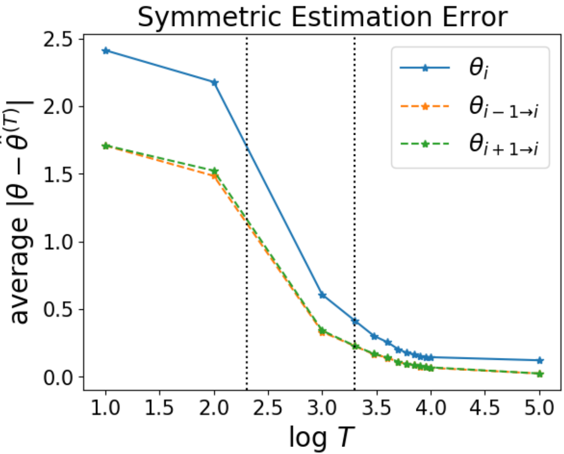


Convergence of Asymmetric Dynamics

- In general, if social biases are asymmetric, unknown whether time dynamics will converge to a stationary distribution

Convergence to Symmetric Equilibrium

- using MCDL, observe convergence to symmetric equilibrium



Allocation Selection Based on Steady-State

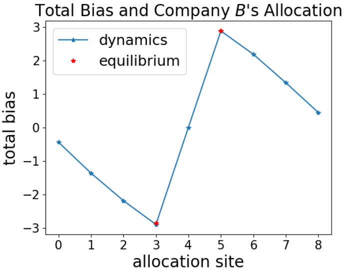
- suppose Company A makes allocation selection with respect to steady-state model:

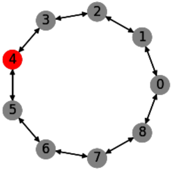
$$Q_{\hat{\mathbf{x}}^{(t_0)}}^{\hat{\theta}^{(t_0)}}(M_A) = \left(\frac{\gamma^{T+1} - 1}{\gamma - 1}\right) \sum_{i \in V} p_i^{(\hat{\theta}^{(t_0)}, M_A)}(\mathbf{x}; \boldsymbol{\theta}) - p_i^{(\hat{\theta}^{(t_0)}, M_A)}(\mathbf{x}; \boldsymbol{\theta})$$

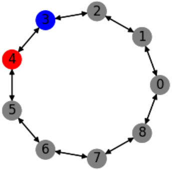
Optimal Allocations Based on Stationary Models

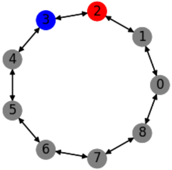
- Company B computes total bias for different allocations

- consumer 3 is optimal
- consumer 5 is worst









Monotonicity of Entropy and Influencing Social Preference

- manifold Gibbs equilibria based on statistic $\boldsymbol{t}$

$$p(\mathbf{x};\boldsymbol{\theta}) = \frac{1}{Z}\exp\Big\{\sum_{i\in V}\theta_i t_i(x_i) + \sum_{\{i,j\}\in E}\theta_{i,j} t_{ij}(x_i x_j)\Big\}$$

- uncertainty due to $\hat{\mathbf{x}}^{(t_0)}$ and $\hat{\theta}^{(t_0)}$
- decreasing entropy corresponds to concentration of preference
- knowledge of how entropy changes with respect to increasing bias parameters can be used in allocation selection

Positive Correlation and Monotonicity of Entropy

- **positive correlation**: Griffiths ['67] showed for Ising model with $t_i(x_i) = x_i$, $t_{ij}(x_i, x_j) = x_i x_j$ and $\boldsymbol{\theta} \succ 0$,

$$\mathbf{cov}\left(t_i(X_i), t_{jk}(X_j, X_k)\right) > 0$$

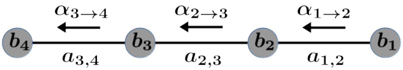
- can show that $H(\mathbf{X}; \boldsymbol{\theta})$ monotone decreasing in $\boldsymbol{\theta}$ for positively correlated $\boldsymbol{t}$ [R and Neuhoff, ISIT '09]

$$\frac{\partial H(\mathbf{X}; \boldsymbol{\theta})}{\partial \theta_i} = -\sum_{j \in V} \theta_j \ \mathbf{cov}\left(t_i, t_j\right) - \sum_{\{k,l\} \in E} \theta_{k,l} \ \mathbf{cov}\left(t_i, t_{kl}\right)$$

Subset Monotonicity In Positively Correlated Ising Trees

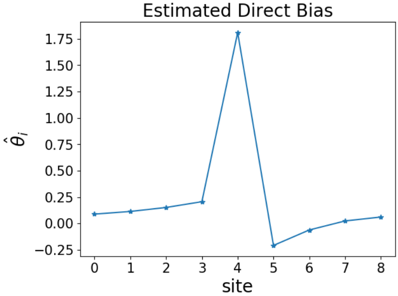
- showed that entropy is monotone decreasing in parameter $\boldsymbol{\theta}$ for an arbitrary subset in family of Ising models on a tree [R and Neuhoff, ISIT 2019]

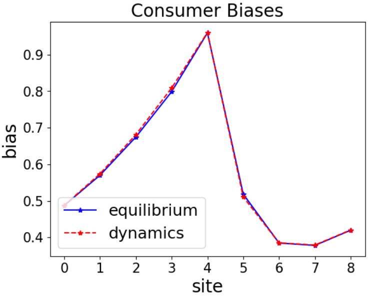
- argued by showing that *messages* used to compute probabilities are monotone in $\boldsymbol{\theta}$



Stationary Equilibrium: Re-Distribution of Direct Biases

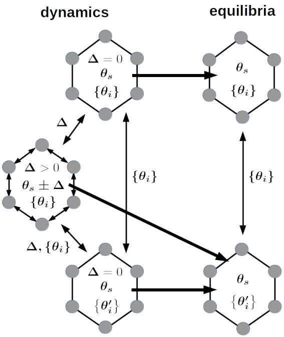
- using MCDL to track direct and social biases, observe a re-distribution of direct biases in the symmetric equilibrium
- note if the direct bias was due to marketing by Company B, re-distributed direct biases would be flipped





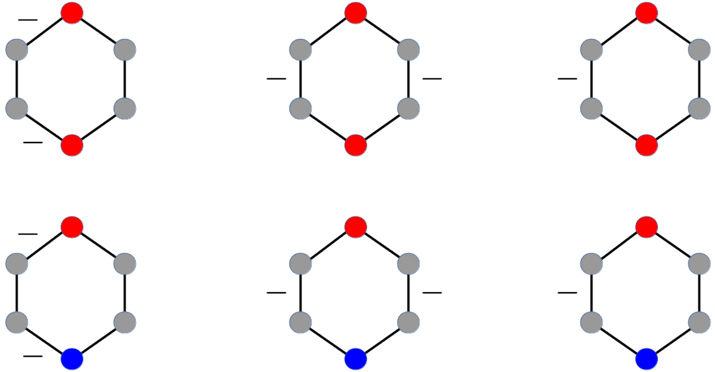
Equivalence Classes of Dynamics Models

- want to understand to what equilibria different dynamics models converge
- want to understand statistical properties of these equilibria that may provide guidance in resource allocation



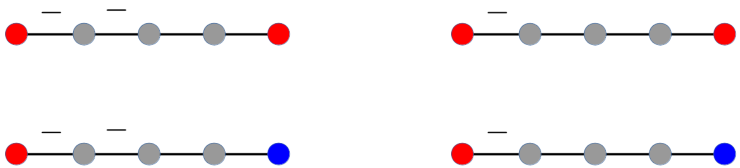
Frustration and Positive Correlation

- if we define *frustration* as the absence of a *ground state*, then preliminary analysis suggests that positive correlation corresponds to non-frustration



Frustration and Positive Correlation

- if we define *frustration* as the absence of a *ground state*, then preliminary analysis suggests that positive correlation corresponds to non-frustration
- frustration can occur in acyclic models
- in traditional statistical mechanics analysis, frustration defined according to parity of anti-coordinating social biases on a cycle



Pulling on this Thread...

- if we can connect patterns of anti-coordinating social biases and direct biases with increasing or decreasing entropy (concentration of choice), can make allocation decisions without computing probabilities

- may be able to avoid computational cost of Monte Carlo simulation

Concluding Remarks

- introduced a model of marketing-influenced consumer decision-making on a social network based on random utility
- machine learning algorithms needed to infer preferences
- model is amenable to reinforcement learning paradigm
- interesting problems in non-equilibrium statistical mechanics
- connecting behavior of entropy with patterns in the polarity of direct and social biases

- Thank you!

- Questions?