

2017 年度科学技術インタープリター養成プログラム修了論文

確証バイアスによるネットワークの分断と凍結：  
Post-truth の理解に向けて

Division and Freeze of Network by Confirmation  
Bias: For Understanding of Post-truth

2018 年 3 月

東京大学 大学院総合文化研究科 広域科学専攻 修士課程

科学技術インタープリター養成プログラム 1 2 期生

大窪 健児

指導教員 松田 恭幸教授

要旨 .....	179
1. 序論 .....	181
1-1. 背景 .....	181
1-2. Post-truth 考察 .....	181
1-3. Post-truth、エコーチェンバー、フィルターバブル .....	182
1-4. 確証バイアス .....	182
2. (方法) 基本的なモデル：空間なし、相互作用数変動 .....	183
2-1. モデルの設定と意味 .....	183
2-2. ネットワーク学習モデル .....	184
2-3. 初期条件 .....	185
2-4. パラメータ .....	185
2-5. ネットワークに対するノイズ .....	185
3. (方法) 2次元モデル：空間あり、相互作用数固定 .....	185
4. (結果) 基本的なモデル：空間なし、相互作用数変動 .....	186
4-1. 意見の変化と凍結 .....	186
4-2. ネットワークの分断 .....	186
4-3. ネットワークの変更を加えた場合 .....	187
4-4. 相互作用数の変動 .....	188
5. (結果) 2次元モデル：空間あり、相互作用数固定 .....	189
5-1. 意見の様子 .....	189
5-2. 特殊エージェントがない場合 .....	190
5-3. 中央に「エビデンス」エージェントがいる場合 .....	190
5-4. 外場（マスコミ）を与えた場合 .....	191
6. 結論・議論と展望 .....	191
謝辞 .....	192
文献 .....	193
インタープリター養成プログラムを受講して .....	194

## 要旨

客観的なエビデンスに基づかない情報によって多数の人間が判断を下してしまう **post-truth** について、数理モデルによる現象の再現を行い、**post-truth** 緩和策の評価やそもそも真実とは何かという問いに迫る。本研究では **post-truth** が生じる前提としてフィルターバブルやエコーチェンバーの生成について考える。数理モデルにより、現実のネットワークに対しては難しい実験が可能になる。

数理モデルの構築にあたっては、エコーチェンバーと深く関係しているとされている確認バイアス（自分の信念に近い情報ばかりを集めてしまうという人間の傾向）に基づき、ネットワーク上の学習モデルを応用することにした。このモデルには、まず各エージェントが相互の意見に対して影響を与えあう議論の段階がある。次に意見が一致するエージェントとのつながりはポジティブな方向に、意見が相違するエージェントとのつながりはネガティブな方向に変化するよう設定した学習の段階がある。

このような数理モデルに基づきシミュレーションを行った結果、ネットワークの分断や意見の凍結が見られた。エコーチェンバー緩和策の一つとして検討されている分断されたネットワーク間を無理やりつなぐという方法では、大量の変更を分断が生じる前に加えた場合に限り、ネットワークの分断が緩和されることがわかった。

エビデンスやマスコミのようなエージェントや状況を仮定し、**post-truth** の生成についても検討を行った。客観性のみを持つエビデンスの存在は意見にさほど大きな影響を与えなかった。一方マスコミは空間的なつながりがある場合に、周囲の意見の変化と同じ方向に意見が変化する様子が見られ、マスコミの存在がある一部の意見を同調させてしまう効果を見出すことができた。

## Abstract

About, post-truth, which can be said as the situation that many people make a decision based on not objective information, mathematical models were made for reproducing post-truth. This model might evaluate the relaxation plan against post-truth or approach the basic question, “What is the truth?” This research focuses especially on generating filter bubble and echo chamber. Theoretical model enables us to conduct some difficult experiment in the real world.

On the construction of the theoretical model, learning theory in networks is applied because it is said that echo chamber arises from confirmation bias, which is the tendency of human that one is more likely to choose the similar information to his/her creed or belief. This model contains two phases. The first one is the phase of discussion which the opinion of each agent influences other’s opinion. Next, the phase of learning that the connection between the agents who have the same/different opinions is made to become positive/negative.

The simulation of the model was conducted and the result suggested the division of network and the freeze of opinions. When the separated interactions are forced to connected, this is one of the scheme for the mitigation of echo chamber, the

network is less divided only if a large amount of change is introduced before the network division.

The occurrence of post-truth is tested with adding an evidence agent or outer influencer as mass media. The existence of evidence agent, who receive no change of network connections, did not affect the freeze of opinions largely. On the other hand, mass media made neighbor change their opinion into the same direction if the distance of space is defined. This means that mass media synchronize the opinions in a part of the population.

## 1. 序論

### 1-1. 背景

Oxford Dictionaries は 2016 年の単語として、“post-truth”を選んだ[1]。Post-truth は一般的には「真実や事実によらない状態」として定義できる。古くは 1991 年の湾岸戦争期から使われていた[1]が、Oxford Dictionaries によると、post-truth という語の使用頻度は 2016 年 5 月ごろを境に爆発的に上昇する[1]。2016 年の 6 月にはイギリスで EU 離脱の是非が問われる国民投票が行われた。EU 離脱を強く支持するイギリス独立党 (UKIP) の党首であった Nigel Farage などが、EU 離脱がイギリスのもたらす利益について、事実によらない情報を提供したとして、批判を浴びることになった[2, 3]。これを筆頭に、イギリス EU 離脱の国民投票以降、post-truth は大きな脚光をあびることになる。

イギリスの EU 離脱に加えて、2016 年にはアメリカで大統領選挙が行われ、Donald Trump 氏が選出されるに至ったが、彼の主張の中にはエビデンスや客観的な知識に基づかないものが散見され、ここでも post-truth が大きく取り沙汰されることになる。具体的には地球温暖化は起きておらず、アメリカの産業を圧迫するためのものだとする Trump 氏の主張に関して、Trump 氏の Twitter では”In the East, it could be the COLDEST New Year’s Eve on record.” ([4]) といった発言がなされ、偏った見方によるデータやエビデンスの解釈が拡散される状況に至っている。

現在、post-truth は主に政治における問題であると認識されることが多いが、すでに地球温暖化などがその矢面に立たされているように、政治的な課題に限らず、科学一般に対して、エビデンスや客観的な知識によらない”post-science”が登場し、相対的な科学の地位が貶められることも十分考えられる状況になってきている。よって、post-truth を深く考察し、そこから普遍的な示唆を抽出することで、今後来るべき”post-science”の時代に科学が対応していくことが求められる。

### 1-2. Post-truth 考察

このような post-truth に対しては、様々な学問分野から考察することができる。まずは、歴史的にこのような事象が起きてきたかどうかを考えることができる。例えば、中世においては現代（あるいは古代ギリシャで）認識されているような科学的知見に反する考え方がヨーロッパにおいて広く流布されていた。その後、そのような権威的な知識に対抗するために近代的な自然科学が誕生していった。現代においてはむしろその科学が支配的・権威的な状況になっており、それに対抗するために post-truth が発生してきたとする見方もできる。次に、歴史に接続して議論すると、歴史修正主義との関連も考慮することができる。歴史修正主義では、証拠によって裏付けされた歴史や歴史観を、証拠によらないもので変更していくことであり、post-truth 的な状況が起きているということが出来る。歴史修正主義が流布する一つの要因として考えられるのは、教科書的な知識が（証拠に基づいて）度々書き換えられ、それが広く報道されることにより、かえって「教科書と違うものが真実である」というような背景を確立してしまうということである。このような歴史教育、歴史とコミュニケーションの問題とも接続して考えられる課題である。アメリカ的に絞った議論をすると、マスメディア自身の偏向報道への批判的意識が、相対的に

Twitter やウェブニュースの権威を高め、post-truth の生まれる原因となっているという指摘もある[5]。Trump 氏の移民政策などとも接続させると、アメリカは「2016年の大統領選において、パラノイア的状况に陥り、対外的に不必要な恐怖心を抱くようになっていた。」[6]という指摘もある。

### 1-3. Post-truth、エコーチェンバー、フィルターバブル

Post-truth 研究において着目しなければならないのは、人と人とのネットワークの状態である。[7]ではアメリカ大統領選における Twitter 上のネットワークを解析し、主要候補であった Trump 氏支持のネットワークと Clinton 氏支持のネットワークで分断が起きていたことを指摘した。ネットワークが分断している状況では、ネットワーク内の情報は頻繁に流通されるが、ネットワーク外の情報はほとんど入ってこなくなる。このような状況では、たとえ誰かがエビデンスや客観的な知識に基づいて批判・検証を行ったとしても、そういった動きが post-truth を信じている人々に届かないことが容易に想像できる。すなわち、このようなネットワークの分断が post-truth が固定化される前提となっていると言えよう。ここで、ネットワークが複数の集団に分断するとともに、ネットワークの内部では内部から得られる情報が支配的になる状況をフィルターバブルと呼ぶことにしよう。フィルターバブルは本来検索エンジンが検索を学習することで、ユーザーが得られる情報が制限されることを意味するが、ここで「検索」を「インターネット上のネットワークを介して移動する情報の受信」を意味するものとしよう。また分断されたネットワーク内で意見が過激化することをエコーチェンバーと呼ぶことにする。ここで考慮したい「過激化」とは、分断されたネットワークの内部において提供される情報が、最初は「意見」だったものが「仮説」になり、さらに「定説」「真実」と変化していくような状況であると読み替えて考える。

このようなネットワークの分断とその内部での意見の過激化は「陰謀論」や「宗教」の問題として古くから扱われており、我々にとっても身近な問題である。これはすなわち、フィルターバブル、特にエコーチェンバーが近年新しく登場した概念であるというよりは、人間と人間が構成する社会において普遍的に存在する傾向であることを示唆している。それでは、その「人間と人間が構成する社会において普遍的に存在する傾向」とは何なのだろうか。本研究はここに問題意識をおくことにする。

Post-truth や「陰謀論」、「宗教」に関しては様々なアプローチ方法があるが、本研究では post-truth などの問題においてもっとも普遍的に存在する目完イズムを抽出することに主眼を置くこととし、数理モデルの構築とシミュレーションによって考察を行うことにする。

### 1-4. 確証バイアス

このような人間社会のネットワークが分断する現象について、[9]などでは確証バイアスによる説明を行っている。確証バイアスとは「自分の信念・信条と一致する情報ばかり集めてしまう人間の傾向」のことを指す。確証バイアスの有名な例として、ウェイソン選択問題がある[10]。ウェイソン選択問題では図1のような4枚のカードについて「偶数の書

かれたカードの裏側は赤である」という命題を検証しようとする。4枚のカードは片面には色、もう片面には数字が示されている。この場合「片面偶数ならばもう片面は赤」という命題であるから、偶数である8のカードの裏が赤であることと、その対偶である「片面赤でないカードの裏は奇数である」を確かめるために茶色のカードをめくることが正しい探索である。しかし、確認バイアスがある状況では、「片面偶数、片面赤」という情報に引きずられることにより、8のカードと赤のカードを確認してしまう。



図1 ウェイソン選択問題で用いられるカード (Wikipedia 日本語版「確認バイアス」より引用)。色については右から順に茶色、赤色のカードが並んでいる。「片面偶数ならばもう片面は赤」という命題であるから、偶数である8のカードの裏が赤であることと茶色のカードをめくることが正しい探索である。しかし、確認バイアスがある状況では、8のカードと赤のカードを確認してしまう。

さて、このような確認バイアスは、我々は自分の信念と似通った情報を集めたがるということを説明する。ということはすなわち、自分の信念と似通った情報源から情報を得る機会が増加し、その情報源とのつながりが強化されるということを示す。このような状況は「自身と同じ意見を持つ人間とのつながりを強化する」というような形で人間社会のネットワークが学習していくものとして説明することもできる。このような考察から、学習理論に近い形のネットワーク学習モデルを構築し、シミュレーションを行うことにした。

## 2. (方法) 基本的なモデル：空間なし、相互作用数変動

### 2-1. モデルの設定と意味

本研究で行なったシミュレーションは、多数のエージェントを用意し、各エージェントが決められたルールに従って行動するというエージェントベースモデルに従っている。また、ネットワークが学習するという点で、ニューラルネットや強化学習、スピングラスとも類似するモデルである。

エージェント  $i$  は  $+1$  または  $-1$  で表された意見  $s_i$  を持つものとする。エージェント  $i$  はネットワークの中で影響を受けるものとする。あるエージェント  $j$  が他のエージェント  $i$  から受ける影響の大きさを  $J_{ij}$  とする。 $J_{ij}$  は整数値で表されるものとする。本研究では、まずは、各エージェントは自分単独で意見を決めるのではなく、他者からの影響の総和で意見を決めるものとする。 $J_{ij}=3$  とすると、エージェント  $j$  が  $-1$  の意見を持っていたとき、エージェント  $i$  が受ける影響は  $J_{ij}s_j=3 \times (-1)$  となる。このモデルでは、全てのエージェントから影響を受けうるものとするので、 $i$  が受ける影響の総和は  $J_{ij}s_j$  を全ての  $j$  について足

し合わせたもの ( $\sum J_{ij}s_i$ ) となる。エージェント  $i$  の意見  $s_i$  は  $\sum J_{ij}s_i$  が正ならば +1、負ならば -1 となる。このような計算によって各エージェントの意見を計算していく。このフェイズを議論のフェイズと呼ぶことにする。議論のフェイズでは各エージェントの意見が時間に伴って変化していくことが考えられる。時間による変化については、変化しなくなるか、十分時間がたつまで計算を行うものとした。

上記のようにして決定された意見をもとにネットワークの変更を行う。エージェント  $i$  と  $j$  の意見  $s_i$  と  $s_j$  が同じ場合は影響  $J_{ij}$  に +1 し、異なる場合は -1 する。 $J_{ij}=0$  の場合は変更しないものとした。これは確証バイアスから与えられるような、意見が同じ者とのつながりを強化し、そうでないものとのつながりを弱める、あるいは反対する効果を示している。この操作を行う段階を学習のフェイズと呼ぶことにする。

上記の議論のフェイズと学習のフェイズを繰り返し行なっていく。

## 2-2. ネットワーク学習モデル

ある  $g$  番目の問題に関してエージェント  $i$  がエージェント  $j$  から受ける相互作用(議論による影響)を  $J_{ij}(g)$  とする。ある問題について議論を行うに当たっての時刻を  $t$  とする。問題  $g$ 、時刻  $t$  でのエージェント  $i$  の意見を  $s_i(g, t)$  とし、+1 と -1 の二値で表されているものとする。 $\text{sgn}$  を階段関数(負の入力には -1、ゼロの入力には 0、正の入力には +1 を返す関数)とすると、相互作用に基づくエージェント意見の更新規則は

$$s_i(g, t+1) = \text{sgn} \left[ \sum_{j=1}^N J_{ij}(g) s_j(g, t) \right]$$

として記述される。この更新規則に基づいて、エージェントの意見に変動がなくなるか、または  $t=1000$  になるまで計算を続けた。

次にエージェント意見の更新が終了した時の意見  $s_i(g, t)$  を用いて相互作用  $J_{ij}(g)$  を変更する。 $J_{ij}(g)$  の変更にあたっては確証バイアスと同様になるように、「意見が同じものからの影響を増加させ、意見が異なるものからの影響を減少させる」ものとした。ただし、 $J_{ij}(g) = 0$  の場合は変更しない。

$$J_{ij}(g+1) = \begin{cases} J_{ij}(g) + 1 & (s_i(g, T) = s_j(g, T) \text{ and } J_{ij}(g) < A \text{ and } J_{ij}(g) \neq 0) \\ J_{ij}(g) - 1 & (s_i(g, T) \neq s_j(g, T) \text{ and } J_{ij}(g) > -A \text{ and } J_{ij}(g) \neq 0) \\ J_{ij}(g) & (\text{otherwise}) \end{cases}$$

この更新規則は微分不可能であり、物理・数学的には不自然ではあるが、シンプルであるため、初歩的な観察には適していると考えている。 $g=100$  になるまでシミュレーションを行った。

このモデルは、意見が同じものとのつながりをポジティブに、意見が異なるものとのつながりをネガティブにするという点で[8]と類似している。しかしながら、本研究で用いているモデルは、議論のフェイズにおける自由度を含めつつ、パラメータを  $A$  の一つに絞っている。このため、最低限の仮定から最低限のモデルを作成しているものと考えている。

### 2-3. 初期条件

0 番目の問題での相互作用  $J_{ij}(0)$  は  $-A$  から  $+A$  までの値を等確率でランダムにて与えた。各問題において時刻 0 でのエージェント意見  $s_i(g, 0)$  は問題ごとに異なる  $-1, 0, +1$  の一様乱数で与えた。

### 2-4. パラメータ

このモデルにはパラメータ  $N$  と  $A$  がある。パラメータ  $N$  はエージェントの数を表す。 $A$  は単純には影響  $J_{ij}(g)$  の最大値と最小値を表すが、このモデルでの  $J_{ij}(g)$  の更新規則は  $+1$  か  $-1$  であったので、 $A$  は  $J_{ij}(g)$  の最大値と最小値だけでなく、1 回の  $J_{ij}(g)$  の更新で  $J_{ij}(g)$  が受ける変更の相対的な大きさを決める量としてみる事ができる。言い換えると、1 つの問題で受ける変更はネットワーク全体に対して  $\pm 1/A$  であることになる。さらに、 $A$  を大きくすることはすなわち  $J_{ij}(g)$  の取りうる値の数が増えることに相当する。これはつまり、パラメータ  $A$  は影響  $J_{ij}(g)$  のバリエーションの大きさを示す量であることになる。

### 2-5. ネットワークに対するノイズ

Post-truth 対策として提案されている [9]、分断したネットワークを強制的に接続する手法の影響を評価するために、 $g=20$  において全く同じ  $J_{ij}(g)$  のコピーを作成し、確率  $C$  でゼロでない相互作用の正負を反転させた。エージェント意見の初期条件  $s_i(g, 0)$  は同一にして変更しないネットワークと変更したネットワーク両方でダイナミクスとネットワークの変更を行っていき、比較した。

## 3. (方法) 2次元モデル：空間あり、相互作用数固定

相互作用数を固定した場合にどのような変化が起きるかを観察するために、影響を受ける相手の数を制限したモデルを作成した。結果の見易さ、相互作用する相手の限定しやすさのために 2次元系にしているが、「相互作用する相手を限定する」という前提を導入していれば、ほかの次元でのモデルと等価である。エージェント意見更新規則の式は  $g$  番目の問題、時刻  $t$  における 2次元サイト  $(i, j)$  におけるエージェントの意見を  $s_{ij}(g, t)$ 、相互作用を  $J_{ijmn}(g)$  として

$$s_{ij}(g, t + 1) = \text{sgn} \left[ \sum_{m=i-B}^{i+B} \sum_{n=j-B}^{j+B} J_{ijmn}(g) s_{mn}(g, t) \right]$$

となる。このモデルでは新たに  $B$  というパラメータが導入された。これは相互作用の範囲を示すパラメータとなる。すなわち、自分の周囲にある一辺  $2B$  の正方形の中にいるエージェントとのみ相互作用する。これを一定ステップ  $T$  繰り返す。その後、影響  $J_{ijmn}(g)$  を

$$J_{ijmn}(g + 1) = \begin{cases} J_{ijmn}(g) + 1 & (s_{ij}(g, T) = s_{mn}(g, T) \text{ and } J_{ijmn}(g) < A) \\ J_{ijmn}(g) - 1 & (s_{ij}(g, T) \neq s_{mn}(g, T) \text{ and } J_{ijmn}(g) > -A) \\ J_{ijmn}(g) & (\text{otherwise}) \end{cases}$$

として  $J_{ijmn}(g)$  を変更する。ここでも、サイト  $(m, n)$  はサイト  $(i, j)$  の周囲にある一辺  $2B$  の正方形の中のサイトに限定されている。相互作用  $J_{ijmn}(g)$  の発展規則は基本モデル (第 2 節) と異なり、 $J_{ijmn}(g)=0$  の場合でも相互作用の変更を行う。シミュレーションは  $100 \times 100$  の 2 次元周期境界条件サイトにおいて行った。初期条件は基本モデル (第 2 節) と同様に与えている。  $g=60$  になるまでシミュレーションを行った。

#### 4. (結果) 基本的なモデル：空間なし、相互作用数変動

##### 4-1. 意見の変化と凍結

各エージェントの意見を図 2 に示した。基本的に意見は時間によって変化しなくなるまで計算しているものとしているから、特に記述しない限り表示するデータは問題番号  $g$  (図中では gener) によるものである。

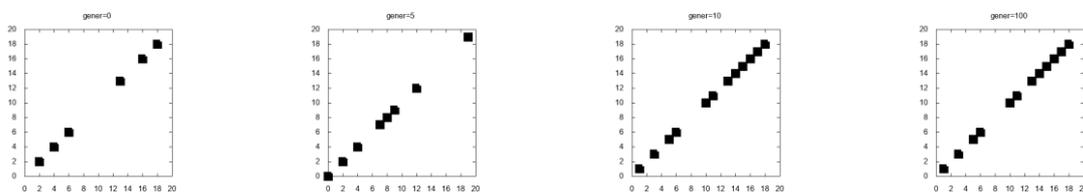


図 2 各エージェントの意見の様子。+1 の意見の場合は白、-1 の意見の場合は黒で表示している。斜めに 20 エージェントの意見をプロットしている。左から  $g=0, 5, 10, 100$  の問題における最終的な意見をプロットした。  $G$  の数が増え、学習が進むにつれて意見の凍結が起きていることがわかる。(  $A=10$  )

各エージェントの意見を問題ごとに見ていくと、学習した問題の数が増えるに従って意見が固定化される様子が見える。このシミュレーションでは、各エージェントの意見の初期値はランダムに与えているから、このような固定化はネットワークの学習によるものである。このようなエージェントの意見の固定化をここでは凍結と呼ぶことにする。

凍結を確認するために、前の問題  $g-1$  と  $g$  の差分を図 3 に示した。学習が進むにつれて凍結が起きていくことがわかる。

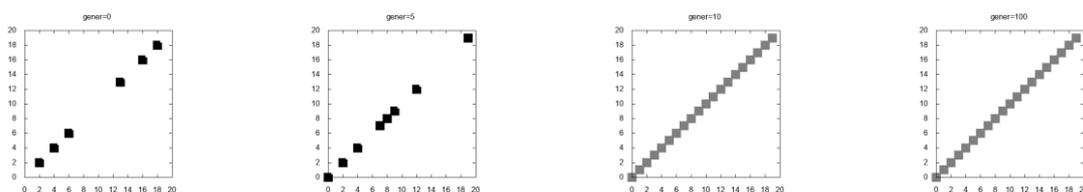


図 3 前の問題  $g-1$  の意見と  $g$  の意見の差分。黒は前の問題と比べて -2、白は +2、灰色は前の問題と比べて変化がないことを示す。問題  $g=10, 100$  では凍結が起きていることがわかる。(  $A=10$  )

##### 4-2. ネットワークの分断

次に、ネットワークを表現する  $J_{ij}(g)$  を取り上げる。  $J_{ij}(g)$  の具体的な値をプロットする

と図4のようになる。学習が進むにつれて  $J_{ij}(g)$  の値が  $-A$ 、 $0$ 、 $+A$  の3つの値に分離していくことがわかる。

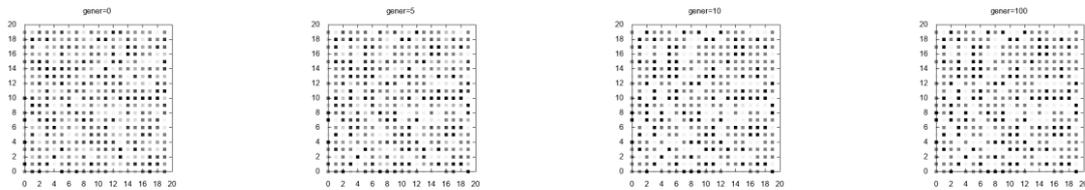


図4 20エージェントと20エージェントの相互作用を記述する行列  $J_{ij}(g)$  の値を色の濃さで表したもの。  $A=10$  であるから、黒が  $-10$ 、白が  $+10$  を表す。縦に配置された0から19までのエージェントが横に配置された0から19までのエージェントに色で示された値の分だけ影響を与えるものとする。

また、ネットワークの構造を  $J_{ij}(g)$  の重みに従って配置すると図5のようになり、大きく2つの集団に分断されていることがわかる。集団の中では  $J_{ij}(g)$  が正となっており、互いに発信した情報から大きな影響を受けていることがわかる。一方で集団の間では  $J_{ij}(g)$  が負となっており、分離した集団の間では反対し合うような影響を与えていることがわかる。

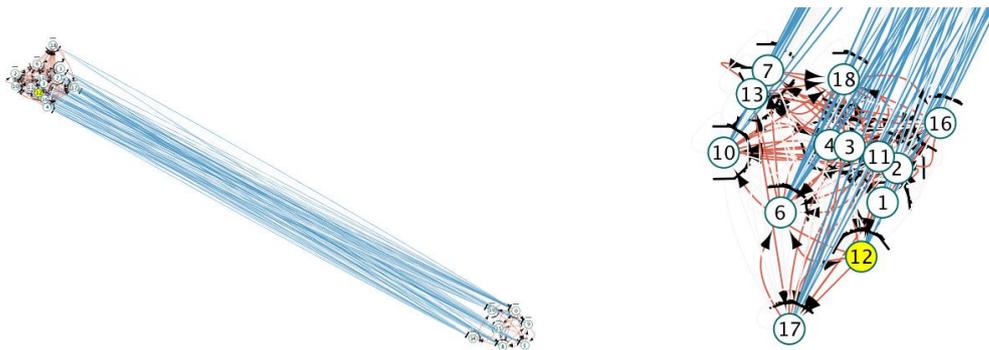


図5 ネットワークを重みづけに従って配置した図。右は分断した片方の集団を拡大したものである。ネットワークが大きく分断していることが確認できる。また、分断されたネットワークの内部では同調する影響（矢印・赤）が多いのに対し、分断された集団の間では反対する影響（T字矢印・青）が支配的になっていることがわかる。

#### 4-3. ネットワークの変更を加えた場合

2-5 で説明したような変更を加えた場合の意見の変化を図6に示す。意見については、変更を加える割合  $\mu$  が  $0.4$  と大きい場合に、変化が起きていることがわかる。

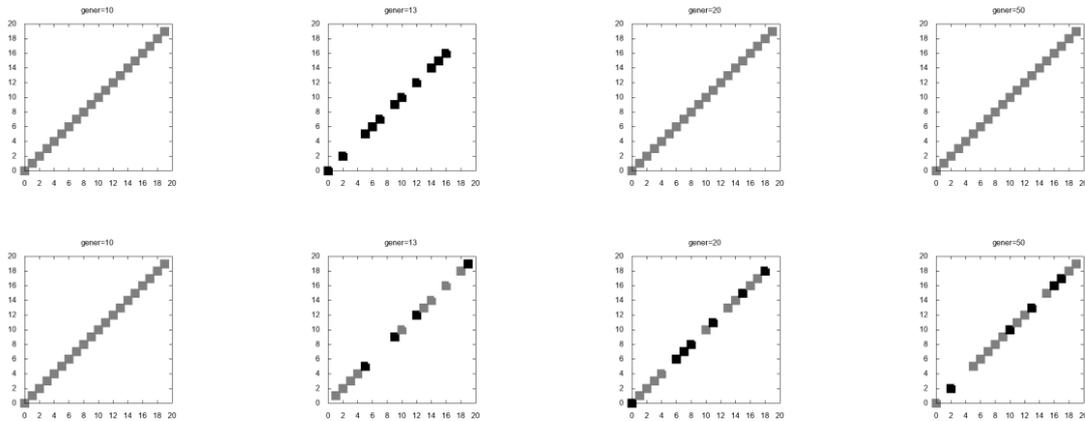


図6 ネットワークに変更を加えた際の意見の差分。ネットワークに変更を加えた場合の意見と加えなかった場合の意見の差を表示している。黒は-2、白は+2、灰色は変化がないことを示している。上段は変更を加えたのが10% ( $\mu=0.1$ ) の場合で下段は40% ( $\mu=0.4$ ) の場合である。左からそれぞれ  $g=10$ 、13、20、50 である。変更は  $g=10$  で加えており、これ以前に変化は見られない。

一方で、この時のネットワーク構造を解析したところ、分断は解消されていないことがわかった。

#### 4-4. 相互作用数の変動

ネットワークの分断を構造や配置とは異なる視点から評価するために、1 エージェントが受ける影響の平均の個数を計算し図7に示した。

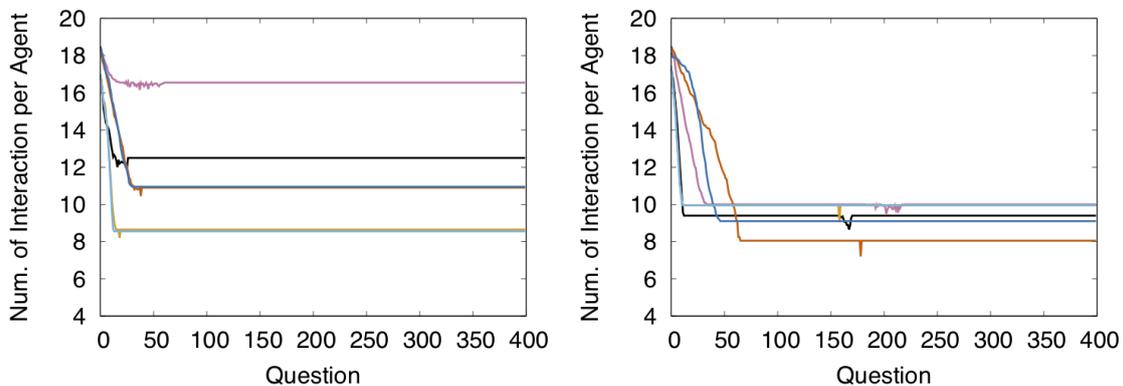


図7 影響を受けるエージェントの個数の平均値。具体的には1 エージェントあたりのゼロでない  $J_{ij}(g)$  の個数の平均である。左が  $g=10$  で返納を加えた場合で、右が  $g=150$  で変更を加えた場合である。色分けはピンクが  $A=30$  かつ  $\mu=0.4$ 、黒が  $A=10$  かつ  $\mu=0.4$ 、赤が  $A=30$  かつ  $\mu=0.1$ 、オレンジが  $A=10$  かつ  $\mu=0.1$ 、青が  $A=30$  かつ  $\mu=0.0$ 、水色が  $A=10$  かつ  $\mu=0.0$  で対応している。基本的には  $A$  の値で大きく変化することが左のグラフから読み取れるが、 $\mu=0.4$  の時は平均個数の現象が大きく抑えられている。しかし一方で  $g=150$  で変更を加えた場合は一時的に平均個数が変動するものの、すぐにまた元の値に戻り、分断の抑制は見られなかった。

ネットワークの変更を  $g=10$  で与えた場合と  $g=150$  で与えた場合を比較すると、 $g=150$  で変更を与えたものは、一時的に相互作用数が減少したものの、すぐに元の値に戻っていることがわかる。それに対して  $g=10$  で与えると、 $A$  の値や加えた変更の割合  $\mu$  などによって相互作用数の収束値が変化することがわかる。特に  $g=10$  で変更を加えて、 $A=30$ 、 $\mu=0.4$  の場合は相互作用数の低下が大きく抑えられていることがわかる。この時のネットワークの構造を配置すると図8のようになり、確かに分断が緩和されていることがわかる。この構造をよく観察すると、確かに緩やかな分断は起きていても、集団内で負の相互作用があり、集団間で正の相互作用がある。ここから分断の緩和を目的としたネットワークの変更はある程度大規模かつ分断が固定化される前に行う必要があると言える。

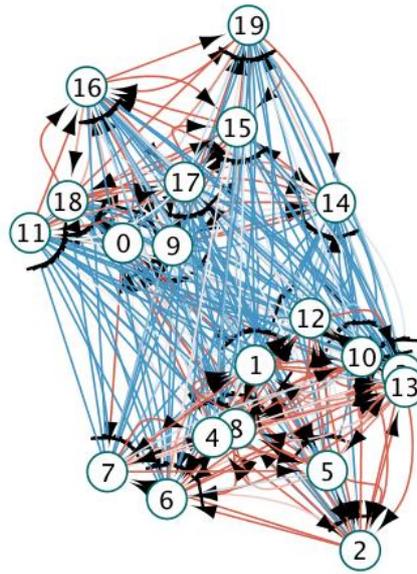


図8  $A=30$ 、 $\mu=0.4$  の場合におけるネットワーク構造を重み付けによって配置した図。緩やかな分断が見られるが、分断した集団内で負の相互作用が見られ、集団間で性の相互作用がある場合があり、分断が緩和されている様子が見られる。

## 5. (結果) 2次元モデル：空間あり、相互作用数固定

### 5-1. 意見の様子

2次元モデルでの意見の様子を図9に示した。基本モデルと同じように凍結が見られるが、2次元モデルで空間を考慮すると、凍結している領域と凍結しない領域が存在することがわかる。このような、凍結する領域のことを凍結相、凍結せずに揺れ動く領域をランダム相と呼ぶことにする。

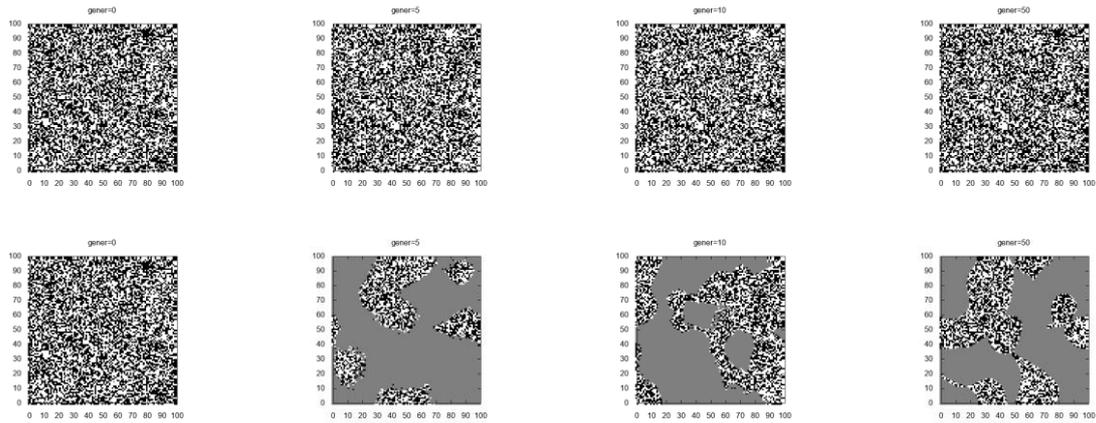


図9 各エージェントの意見の様子。縦に100エージェント、横に100エージェントで合計1万エージェントについて計算している。上段は意見をそのまま図示しているが、情報を抜き取りづらいため、下段に前の問題  $g-1$  の意見と  $g$  の意見の差分を示した。ランダムに変化するランダム相と前の問題と同じような意見を持ち続ける凍結相が見られた。問題番号は左から  $g=0, 5, 10, 50$  である。 ( $A=10, b=10$ )

### 5-2. 特殊エージェントがない場合

$A = 10$  で固定し、影響を受けるエージェント数のパラメータ  $b$  を変化させてシミュレーションを行った。パラメータ  $b$  が小さいとランダム相と凍結層が世代ごとにランダムに発生する様子が見られる。パラメータ  $b$  が中間領域になると、稀に大きな範囲でランダム層が発生することがある。これは、全体を覆う凍結相が周期境界条件によってフラストレーションを起こし発生したものと考えられる。さらに  $b$  を大きくするとエージェントの意見は完全に凍結する。

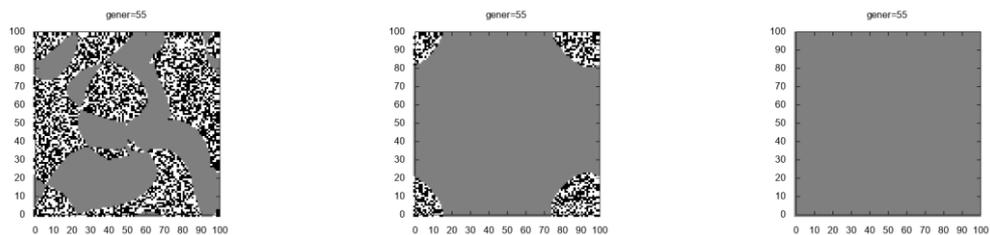


図10  $b$  を変化させていったときの凍結相とランダム層の様子。全て  $g=55$  の時の  $g=54$  との意見の差分を示している。左から  $b=10, 25, 45$  である。パラメータ  $b=10$  の時はランダム相と凍結相が複雑に分布している様子が見られる。 $b=25$  の時には、円形の領域でランダム層が発生している。 $b=45$  になると、全体が凍結層となる。

### 5-3. 中央に「エビデンス」エージェントがいる場合

2次元サイト中央に全てのエージェントと結合し、かつ他のエージェントから影響を受けない ( $b = 50, J_{ijmn}(g+1) = J_{ijmn}(g)$ ) エージェントを設置したが、大きな変化は見られな

かった。

#### 5-4. 外場（マスコミ）を与えた場合

外場  $h$  を与えた。これにより議論フェイズにおける更新規則は

$$s_{ij}(g, t + 1) = \text{sgn} \left[ \sum_{m=i-b}^{i+b} \sum_{n=j-b}^{j+b} J_{ijmn}(g) s_{mn}(g, t) + h \right]$$

として変更される。

ランダム相と凍結相の発生についてはやや、ランダム相が生じづらくなっているように見えるが、特殊エージェントがない場合と大差なかった。しかし、ランダム相内では同じ方向に意見を変化させている状況が見られた（図11）。

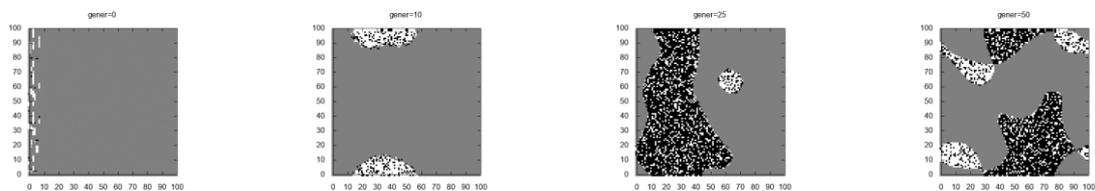


図11 マスコミとしての外場を加えた場合の凍結相とランダム相（ $A=30$ 、 $b=10$ 、 $h=0.001$ ）。ランダム相の生じ方については外場がない場合と大差ないが、ランダム層の内部ではそのうちの多くが一斉に意見を変化させていることがわかる。

## 6. 結論・議論と展望

以上の結果では、確証バイアスを仮定したエージェントにおいて、ネットワークの分断や意見の固定化・先鋭化、一部の集団におけるマスメディアに対しての画一的な挙動が現れることを見出した。このような状況はフィルターバブルやエコーチェンバーの登場を説明するものであるとともに、一部 post-truth の再現になっているとも言える。

また、post-truth 緩和策として提案されている異なる集団間でのネットワークの接続や集団間でのネットワークの変更などを試みたところ、ネットワークが分断する前に変更を加えるとともに、エージェントの影響の受け方のバリエーションが大きい場合にのみ分断が緩和されることがわかった。現実のネットワークでは、今回行なったシミュレーションのように一方向に遷移するようなものではないことが予想されるため、ネットワークが分断する前に変更を加えるというのはむしろ現実のネットワークのほうが容易である可能性はある。例えば、現実のネットワークでは常に新規参入や脱退が存在するが、そのようなエージェントの出入りは部分的にネットワークを初期化しているといえ、分断までの時間を延長する効果があるものと予想される。一方エージェントの影響の受け方のバリエーションについては、「他者とのつながり方は極端なものではなく、グラデーションを持たせた方がネットワークの分断を緩和させる」という示唆を得ることができた。これは、異なる集団間でいわば交渉役や外交官的な存在となるエージェントが存在し得ることを示唆する。

このようなエージェントが存在するのかどうかについてもぜひ調査してみたいと考えている。

本研究の後半では、マスコミやエビデンスを仮定したエージェントや変更を想定してシミュレーションを行い、その結果を考察した。このような2次元系での考察は[8]でも触れられていなかったような効果であると考えている。その過程で、浮かび上がってくるのは、そもそも我々は何を真実としているのか、どのような条件を満たせば真実なのかという哲学的な問いである。Post-truth 研究には常に truth とは何かを対照的に比較しながら考えてゆかなくてはならない。今回は「他者から影響を受けない」という客観性に着目してモデルを構築したが、そう出ない場合と大差は見られなかった。例えば、時間によっても変化せず、常に同じような意見を持っているものとして真実を導入するのもありえるかもしれない。また一方で、真実と言えるものが存在しない課題ももちろん存在する。本研究で得られた結果はそのような真実がない場合にも適応可能であるが、現実の問題として、真実がはっきりしない場合に post-truth と呼ばれるようになったとした場合に我々はどのような行動を取るべきなのか。Post-truth 周辺の課題は山積している。

さらにネットワークの分断一般に目を向けると、[11]のように、現在までに数多く調べられてきた、社会学、心理学での調査をもとに、確証バイアスやインターネットのプロトコル自体がネットワークを分断させる方向に作用することが示唆されている。[11]ではさらにラジオやテレビの普及についてもネットワークの分断が生じることを指摘した。このように普遍的に生じうるネットワークの分断がインターネットの登場によってさらに加速されることについても触れられていたが、本研究における2次元系の結果は、インターネットの登場における「やりとりする人間・情報の数が増える」という点について、ネットワークの分断・凍結が加速されることを指摘している。[11]では他にも、分断れたネットワークでの意見の先鋭化やリンクのダイナミクスについても触れているが、このような指摘との対応は本研究における今後の展望としたい。

## 謝辞

まずは、研究をご指導くださいました、松田恭幸教授に厚く御礼申し上げます。研究の進捗が好ましくないときも様々な話題を提供してくださり、モチベーションの維持に不可欠でした。また、修了研究ゼミやその他の発表会で様々なご指摘をくださった江間有沙特任講師、見上公一特任講師にも深く感謝しています。指導教員と学生ともに物理系出身でしたので、お二方のご指摘はどれも新鮮かつ的確で、インタープリターの研究としての方向づけにはなくてはならない存在でした。廣野喜幸教授と田中幹人准教授には修了研究発表かにおいて大変重要な文献を紹介していただきました。本論文の執筆には不可避な内容でした。この場をお借りして御礼申し上げます。最後に、インタープリターの学生、特に12期生のみなさんにはモチベーション面で大変勇気をもらいました。この研究がここまでできたのも切磋琢磨できる仲間がいてこそだったと思います。ありがとうございました。

## 文献

- [1] Oxford Dictionaries. Word of the Year 2016. URL:  
<https://en.oxforddictionaries.com/word-of-the-year/word-of-the-year-2016>  
2018/2/25 閲覧
- [2] 毎日新聞社 「公約「うそ」認める幹部 「投票後悔」の声も」 URL:  
<http://mainichi.jp/articles/20160627/k00/00e/030/145000c>  
2016年6月27日 2018/2/25 閲覧
- [3] ジョシュ・ロウ ニューズウィーク日本版 「英キャメロン首相「EU離脱派6つのウソ」」 URL: [https://www.newsweekjapan.jp/stories/world/2016/06/post-5289\\_1.php](https://www.newsweekjapan.jp/stories/world/2016/06/post-5289_1.php)  
2018/2/25 閲覧
- [4] Donald Trump Twitter (@realDonaldTrump) URL:  
<https://twitter.com/realDonaldTrump/status/946531657229701120> 2018/2/25 閲覧
- [5] 池田純一 2017年 「〈ポスト・トゥルース〉アメリカの誕生 ―ウェブにハックされた大統領選」 青土社
- [6] Hofstadter 1963 from Imbo 2017 lecture
- [7] Yochai Benkler, Robert Faris, Hal Roberts, Ethan Zuckerman. 2017. Columbia Journalism Review.
- [8] Michela Del Vicario, Antonio Scala, Guido Caldarelli, H. Eugene Stanley & Walter Quattrociocchi. 2017 Modeling confirmation bias and polarization. Scientific Reports
- [9] 長倉克枝 2017年7月 「特集：トランプ VS 科学ネットで軽くなる「事実」の重み」 日経サイエンス2017年7月号 日経サイエンス社
- [10] Foss, Brian M., ed. New horizons in psychology, edited by Brian M. Foss. Baltimore, Penguin Books 1966
- [11] キャス・サンスティーン 石川幸憲訳 2003年 インターネットは民主主義の敵か 毎日新聞社

## インタープリター養成プログラムを受講して

インタープリター講座の面接のことをふと思い出す事があります。当時の研究内容や経歴について様々な質問を受けたわけですが、自分が志望動機欄で仕切りに強調した「議論」について「議論とは何か」「議論で相手の意見を変えることは必要か」という問いを投げかけられました。私のインタープリター講座はそこから始まりました。

様々な授業を受講して、議論とは何かを探してきましたが、科学技術コミュニケーションの周囲に集まる人たちは、科学にすでに興味を持っている人たちや私たちとすでにつながりのある人たちでした。他の受講生の修了研究でも度々話題になったように、科学コミュニケーションの重要な対象である「そもそも興味やつながりがない層」との相互作用はなかなか実現できないというのが実感でした。

むしろそのような問題意識がこの修了研究を突き動かす原動力となっていたようにも思います。すなわち、「なぜ科学に興味のある人と科学に興味のない人で分断が起きるのか」「科学に興味のある人同士でしか議論できないのか」という問いです。この修了研究はモデル設定から解析、結果の解釈まで研究と呼ぶには程遠い出来ではありますが、そこから得られるわずかな教訓を汲むとするならば、「我々はその分断を解消するためにわずかでも不断の努力をし続けなくてはならない」ということが言えるかと思います。

いまでも本専攻の研究室では「インタープリターとは何か」という質問を度々受けます。自分の持つ「インタープリター像」はつかめてきたように思えても、一般的に「インタープリターとは何か」と問われるとどこか散漫で説明しづらいところがあるように思います。しかしながら、この講座を受講して確信を持って言えるのは「インタープリターのいる場所」を私は知ることができたということです。

色々なことは言えるのですが、結局は「インタープリターのいる場所」は12期生の間の中にあるのかもしれませんが。