

不特定多数の空間における言説の安定性に関する理論

Tea partners, Co., Ltd.
CEO Toru Nomakuchi

2015/09/28

1 背景

昨今、SNS等の普及により、不特定多数が多くの意見を交わす機会が増え、それにより知の共有が進むなど、ネットは情報共有の場として存在価値を増してきているといえる。その一方で、建設的な会話とならず、組織的な煽動や、感情的な無意味な罵りあいのような、言説も増大している。俗に言う、荒らしや煽動のような書き込みが増えると、従来の冷静な論者や、一般の論者が沈黙をするようになり、その場全体が荒れて、議論などが不可能になり使い物にならなくなる。

作者もこのような荒らしや煽動をされた経験があるが、この手の荒らしや煽動は、最も論理的な用語に対して、最も感情的な用語（例：ルサンチマン、アスペ、承認欲求、ネトウヨ）でレッテル貼りをすることで、論理的な話を封じ込めようとする手口で、しばし、その場における最も中立的な話者を狙い打ちにしていく特徴がある。このような荒らしや煽動行為がなされる目的は多種多様であると想像されるが、このような煽動行為が場全体を不穏なものとし、多くの煽動家が暗躍する土壌を醸成してしまうことにより、個人攻撃としてのデマの拡散など、いわゆる炎上と呼ばれる現象を引き起こしたりしている。

このような現象を野放しにしていると、ネット全体が不穏なものとなる上、ネットメディアの影響力は非常に大きなものとなっていることや、この煽動から炎上までのメカニズムがファシズムに似ていることから、舵取りを間違えると、社会全体がファシズムに向かうような危険性も孕んでいると考えている。本稿は、場の質を守るために、このような荒らし行為に対抗する手段に関する理論を示すことを目的とする。また、ファシズムに繋がるような危険な言説をどのように封じ込める手法に関する理論を示すことをより大きな目的とする。

2 token 基底変換による感情表現空間の制限

2.1 場における感情表現の尺度である MK 比率

まず、場について定義しよう。冷静な場とは、文脈が存在する場において、事実や共通認識などに基づき、その文脈で論理的な会話を行うことであろう。また、「何が好き」「何が嫌い」といった発言も、例えば、マーケティング戦略を考える際に、その商品が売れるかどうかを考える際の判断基準の範疇の発言としては、統計データの提供という意味合いで論理的なものであるといえるため、この冷静な場における「論理的」とは、ある文脈の議論をする際の判断材料に繋がる論理構成要素と定義できよう。このような言説を場の「主要な」(Major)の意味で、「M 言説」(m remark)と定義することにする。

では、不適切な発言はどのようなものであろうか。「M 言説」の定義にあるように、感情の表明そのものが場を乱すということでもなく、場の文脈とは無関係な感情表現が不適切であると言える。対人的な関係において、感情表現が向けられた場合は、個人攻撃に繋がることもあり、荒らしや個人攻撃に繋がることもあり注意する必要があるだろう。ただ、その感情表現が互いに関わりあいたい相手の判断基準となることもあるため、単純に対人の距離感をとるための範疇においては、感情表現が問題があるとは言い切れないであろう。これは、逆の見方をすると、自発的に距離をとろうとしている相手に対する感情表現や、本人に直接的に行わない陰口のような感情表現は、距離をとるための対人の感情表現ではないため、個人攻撃の嫌がらせ目的の犯罪行為であるといえよう。このような観点から整理すると、その感情表現に目的が無いか、あるいは、その目的が場において相応しくないということが、荒らしや個人攻撃に結びつく言説であるといえる。このような言説を蹴る (kick)、叩く (knock)、死なせる (kill) などの頭文字をとって「K 言説」(k remark) と定義することにしよう。

次に、荒らしや煽動の定義を行う。荒らしや煽動は、「M 言説」が極端に少ないか、「M 言説」を曲解し、「K 言説」に誘導する行為であるといえよう。本稿では、「M 言説」の量が少なく、「K 言説」が多い状態を表すメトリクスとして下記のように「MK 比率」という値を定義し、「MK 比率」が極めて大きい人を「荒らし」と定義し、「M 言説」を「K 言説」に誘導する人を「扇動者」と定義する。

まとめると、適切な場とは、文脈が存在し、「MK 比率」が低く抑えられている場であり、荒らしは「MK 比率」が極めて大きな人であり、扇動者は「M 言説」を「K 言説」に誘導する基底変換をする人と定義できよう。

2.2 場における token 基底変換

言説は多数の sentence、sentence は多数の token から構成されるが、本稿ではこの token に着目して、場における言説について分析する。

まず、荒らしのメカニズムに関して「M 言説」と「K 言説」を用いて説明する。一般に、「M 言説」に用いられる token に対する感情的なレッテル貼りは、「M 言説」を阻害するものである。例えば、エントロピーに関する話をする場において、エントロピーは似非科学などというレッテル貼りをするような煽動があれば、「M 言説」であるエントロピーという token に対して、「K 言説」であるとレッテルを貼るものであるが、そのレッテルの方が正確には「K 言説」であるため、そのようなレッテル貼りをする人を排除することや、「K 言説」であるレッテル貼り用語を排除もしくは無効化することが必要となる。

このような煽動を token 基底空間から見ると、最も論理的な token の「MK 比率」を高く誘導して、関連する論理的な token の「MK 比率」も同様に高く誘導する行為であるといえる。この荒らし行為をされると、「M 言説」に関連する全ての「MK 比率」が高く誘導されるため、場における論理構成が困難になる。このような「M 言説」の論理構築を困難なものとする「M 言説」に対する破壊的な token 基底変換を「破壊的ミーム」と定義する。このような「破壊的ミーム」により煽動された場においては、建設的・論理的な議論が困難となる。何も対策をしなければ、場の参加者の多くが元の文脈を維持する強い意思を持っていなければ、場は感情的な token で埋め尽くされ、「MK 比率」が跳ね上がることになる。

また、このような煽動に容易に煽動されて、「K 言説」を連発するような人が多い場合は、容易に荒れるため、継続性がなくなる。場の「MK 比率」を低く抑えるためには、感情的な表現を抑えるよう token 基底空間を設定する必要がある。

すなわち、token 基底空間からみて、安定的な冷静な場とは、1) 「K 言説」の抑制、2) 「M 言説」の構成要素の token に対して「破壊的ミーム」を埋め込む、感情表現空間を拡大させる基底変換の抑制および無効化、が可能な場ということになる。

2.3 破壊的ミームに対するワクチン

「破壊的ミーム」を組み込まれた、感情的な表現が支配的となってしまった場においては、より刺激的な感情的なレッテル貼り用語が極めて伝搬性の高い token となる。また、そのような感情的な token は感情的な場においては感染性が高く、多くの参加者の思考回路にそのような感情的な token が入り込んでくることになる。逆に考えると、最も感情的な token を最も論理的な token と解釈できるように、感情的な token の再定義を行うことで、感情的な表現を抑えこむことができることになる。

例えば、しばし罵倒・侮蔑用語として用いられる「馬鹿」という用語も、語源を辿ると「馬を鹿と認識する人」のことであり、「誤認」のことを意味する極めて論理的な用語であり、その言葉自体にはもともとは侮蔑の意味がない。つまり、場の文脈に沿わず事実認識から外れた発言のことを「馬鹿な発言」と定義すると、「M 言説」の人を馬鹿という人が「馬鹿」という論理となる。すなわち、「M 言説」を「K 言説」だとレッテルを貼る行為が「K 言説」であるという論理に誘導されるため、破壊的ミームを抑制することができる。このような再定義を破壊的ミームに対する「ワクチン」と定義する。

2.4 token 基底変換型のワクチン

ワクチンを体系的に考えた時、token 列による感情的な token の再定義の抑制および無効化ということになるが、これは換言すると、token 基底変換による感情表現空間の制限ということになる。つまり、予め感情的な「K 言説」に繋がるような token を使用不可にするか、対応するワクチンを定義しておき、その定義を適切なタイミングで周知させるメカニズムを用意しておくことが場の安定性を担保するうえで有効な対応策であると考えられる。

3 token 列の学習による感情の誘導

3.1 MK 比率の発散系

時間軸において「MK 比率」が増大し発散している人に対する処置が必要になるケースがある。この場合、確信犯の荒らし、煽動であるため、システムを超えて物理的な実力行使に出てくる危険性すらあるため、そのような行為に走らないよう予防的措置をとる必要がある。

3.2 文脈依存型のワクチン

例えば、不法行為を行った者がいて、犯罪行為を追及した結果、不法行為を行った者が発見者・追及者の殺害を仄めかしを繰り返すような、極めて興奮状態にあるような場合において有効なのが文脈依存型ワクチンである。

その不法行為を行った者が「殺害」について考えた時に、そのような危害を加える行為に及んだ場合に、死ぬのは相手ではなく、脅迫している人が社会的に「死ぬ」（＝この場合、警察に捕ま

るということを表す) というように、これは、自身が相手に加えようとした危害がそのまま自身に戻ってくるという暗示をかけることで、相手の過激な行為を抑制するタイプのワクチンである。

この場合、「殺す」→「(相手が)死ぬ」→「(自分が警察に〇〇の法律で捕まって)社会的に死ぬ」という文脈の連想をしやすいように、「死」という token と「逮捕」という token を強く結びつけることで効果を発揮する。これは、相手の過激な発言にあわせて、繰り返しこの文脈を、ニューラルネットワークの教師信号のように作用させる手法である。相手が「死」について考えたときに、実力行使について考える確率より、逮捕される心配を考える確率が増すまで繰り返すことで、過激な実力行使を抑制するフィードバックループを形成することにより、実力行使することを考えたくなくなるよう誘導するものである。

4 結論

本稿では、token 基底変換型ワクチンと、文脈依存型ワクチンを提案した。token 基底変換型ワクチンは、場の口コミに関してモデル化したワクチンであり、口コミ伝播しやすい感情的な token に対して効果を発揮するワクチンである。一方、文脈依存型ワクチンは、ニューラルネットワーク、ディープラーニングに基づきモデル化したワクチンであり、極めて興奮状態にある人の頭の中で繰り返し出てくるであろう過激な感情的な token に対するワクチンである。本稿で示したこれらワクチンによる対応策は、ある場における、荒らしや煽動者に対して、2年半ほどかけて様々な形で実施したものであり、「ワクチン」は、荒らしを抑制するのに対して有効な手法であることは確認されている。しかしながら、荒らしのパターンは非常に多岐に渡り、次から次へと悪意をもって荒らしや煽動をする人が多く存在する場においては、非常に時間がかかる対処法でもある。また、各「ワクチン」の浸透には時間がかかるため、この手法は、悪意のある荒らしや煽動者を場から排除する方法と併用することで始めて実用的なものとなると考えている。