

論文特集 「Massive Data Flow : 自然や人工のシステムに みられる複雑さを理解する」にあたって

池上 高志
(東京大学)

岡 瑞起
(筑波大学)

栗原 聡
(電気通信大学)

1. Massive Data Flow 分野の立上げ

Massive Data Flow (MDF) とは、爆発的な情報の流れの中で生まれる「自己組織的なパターン」の中に見いだされる「生命性」のことである。同じ巨大データでも、Big Data がデータの量・質的に巨大であることを謳うのに対し、MDF は動きのある巨大なデータのダイナミクスがつくり出す「生命性」に注目する。

2015年、このMDFの生命性は私達の社会の至るところに見ることができる。携帯、ATM、インターネット、交通機関、各種サービス、社会のインフラは知らない間に「人工生命化」していたのである。人工生命化とは、いろいろな技術やサービス、メディアが自動化しその仕組みがどんどん私達からは見えなくなり、私達の制御を離れていき、利便性のためだけではなくシステムのもつ自律性・ホメオスタシス・自己発展性、といった生命的な性質を顕著にもつシステムのことをいう。そうした意味での人工生命化したシステムは、人の時間と空間とは異なるスケールに存在し、作動するものである。例えば、インターネットは、24時間365日動き続け、その帯域幅は人間同士がつくり出せるコミュニケーションを時間スケールでも空間スケールでもはるかに凌駕する。それをマッシブデータフローというわけだが、データフローのもつ情報の中味と同時に、その情報をどう提示するか、どういうタイミングで与えるか、という「メタな情報」が並列的に現れ、それこそが生命性をかたちづくる。システムは「固定された」システムではなく、つねに「動き進化する」システムである。例えばインターネットのそうしたダイナミックな側面を、これまでのセンサネットワーク、ロボット工学、スマートルーム、バーチャルリアリティ、などそれぞれの分野に見いだすことで、新しい科学領域をつくるための礎にしようということからMDFの研究は始まった。最初のオーガナイズドセッション(2011年度人工知能学会全国大会(第25回, 岩手))では、そうした各分野の重鎮と若手の研究者を招待して議論を行った。

2. 生成系と解析系

2011年の全国大会報告は、2012年の本誌特集号(Vol. 27, No. 6, pp. 626-679)にある。このときの注目点の一つが、研究のアプローチを生成系と解析系に分けて考え

たことだ。生成系とは典型的にマッシブデータフローをつくり出す(しかしtoy modelがつくり出すデータの世界は小さい)と同時に、それを制御する機構とともに世界全体を「つくってわかる」アプローチであり、ロボットや人工生命一般のアプローチを指す。一方で解析系は、得られた莫大なデータを、機械学習やさまざまな方法で実社会・世界がつくるマッシブデータフローを理解するために解析・分解・可視化してみせるアプローチである。2011年当時、生成系として石黒浩(大阪大学)、角康之(公立ほこだて未来大学)、廣瀬通孝(東京大学)、前田太郎(大阪大学)、徳田英之(慶應義塾大学)、解析系として笹原和俊(名古屋大学)、橋本康弘(筑波大学)、増田直紀(Bristol大学)、森純一郎(東京大学)(敬称略)らをお呼びした。

この2011年のキックオフでは、生成系の派手なデモンストレーションが目についた。実際、これまで人工生命を牽引してきたのは生成系のアプローチである。それに比べ、解析系のアプローチは何か無理矢理に定量化・視覚化しても、という気がしないでもなかった。しかし5年後の今、人々が注目するのは圧倒的に解析系のアプローチである。G. HintonのDeep Learning、岡田真人らによる圧縮センシングの手法は、ビックデータの分野に大きなブレイクスルーをつくり出した。データ駆動型の知能、学習、それに付随する予期せぬ予測。それらはあつという間にさまざまな分野に拡散していった。逆に生成的なアプローチ、特にこれまで人工生命の研究を牽引してきた「小さいモデル(toy model)」のアプローチは、劣勢に立たされている。その理由の一つは、私達の認識する世界の複雑さが更新されたことにある。それは世界が人工生命化したことの証拠でもある。

3. MDF 2015

2015年度人工知能学会全国大会における5年目のMDFのオーガナイズドセッションでは、岡田真人(東京大学)と津田一郎(北海道大学)を招き、前者は解析系のアプローチから、後者は生成系のアプローチから議論していただいた。今回の論文特集でも、生成系と解析系のアプローチが並立する。

竹内、廣瀬らは、「Human-in the loop」という概念により、人とWebのサービスを有機的に結び付ける。す

なわち、Web上のサービスが自動的に人の意図とは関係なく、あるパターンをつくり出し、人はそれに反応することで行為決定していく。このアプローチは生成系のアプローチが同時に解析的でもあり、人へのフィードバックをさせることで、半自動化された柔らかい制御を実現し、人は本来的にサイバネティックになることで「新しい人」になる。その形を模索しているといえる。ほかにも生成系のアプローチとして中井、筑紫らのアプローチは、GAを用いたソフトウェアのモジュール化、クラスタリングを行いながら、ソフトウェアを改良していくという議論を展開している。どちらも人とソフトウェアのフィードバックをうまくシステムの中に埋め込んで情報の流れを制御する仕組み、と考えることができる。

一方、池上、岡、橋本らは、Twitterの長時間の時系列を用いて、Webの興奮性媒質としての性格、Webの中の情報の流れを解析している。またWebを下層のダイナミクス（パケットの交換）を調べてシステムのロバストネスを議論し、Twitterの解析と合わせてWebに潜む生命性を議論している。一方、笹原らはTwitterの時系列データに潜む「意味的構造」を可視化するための新しい方法を提案している。これらは解析的なアプローチであるが、最初に解析したいものが決まっているわけではなく、巨大なデータの中から発見し、その現象論を編んでいくという意味では、これまでの解析的なアプロー

チをより発見法的な領域に拡張した、といえる。

5年目のセッションとわずか4本ではあるが、この特集で見えてくるのは、人を念頭に置いたWebの生成系のアプローチと、データ駆動がもたらす解析手法の進化である。もちろんどちらが優位ということではない。むしろ、解析的アプローチがデータ収集と解析の自動化を欲し、生成系が観測・行為主体である人間を中に含み、生成系と解析系の境界はますます分ち難くなると思われる。このことはつまり、これまでの技術と異なり、対象とそれを調べる人という分け方ができなくなり、程度の違いはあれすべて動くことで状態が決まり、先に状態を与えられないという意味では、生命的である。社会は人工生命化の方向に加速されるといえる。

4. 結 語

この特集は、そうした生成系と解析系の新しい融合に向けてのしかし極めて過渡的な様相の濃いものである。これからの数年、この融合は急速に進んで新しい研究領域も生まれるに違いない。今後のMDF研究は、人工生命化する社会が、その中で生じる複雑さ（新たな倫理的、社会科学的、言語的、民族的問題も含め）をいかに手なずけることができるか（taming）、ということに向かうだろう。