

社会科学とポリエージェントシステム

出口 弘

1. はじめに

本稿では社会科学にとってのポリエージェントシステム観の意味を説明する。ここで我々がポリエージェントシステム（多主体複雑系）と言っているのは、多くの自律的な主体からなるシステムのことである。自律的な主体からなるシステムについては、すでにマルチエージェントシステムという言い方があるが、我々があえてポリという用語を使ったのは、同質のエージェントの自律分散的活動のみならず、多数の機能分化したヘテロな機能を持ったエージェントからなるシステムの分析を目指しているからである。そこでは、多数のエージェントが集まって（重合≡ポリ）、上位のエージェントを形成するメカニズムやエージェント間の階層的関係もまた分析の対象となる。

このような社会システムや経営組織、経済システムを念頭に入れた多主体複雑系としてのエージェントの科学は確立された領域ではない。我々のポリエージェントシステムという概念自体も、明確な数理的ハードコアを備えたリサーチプログラムではない。むしろ緩いリサーチイニシアチブとでも言うべきものである。

それにもかかわらずこの領域のリサーチプログラムを策定することは、社会科学にとっては長期的に非常に重要な意義を持つと考えられる。そこで本稿では、この自律的エージェントのリサーチイニシアチブとしてのポリエージェントシステムの分析と社会科学の関係について簡単な解説を試みたい。そのために、我々は2つの軸を重視する。1つは複雑なシステムとしての社会科学の特質であり、もう1つは複数のリアリティを扱う科学としての社会学の特質である。

2. 主体を含む複雑系の科学としての社会科学

社会科学はそもそも複数のリアリティを扱う科学である。エージェントが自ら構成したリアリティを扱う科学であると言ってもよい。社会や経済の諸システムでは、そこで登場する、個人や企業、組織などのエージェントが独自のリアリティ（モデル）を自ら構成し、そのモデルに従って活動している。しかし社会科学の多くの領域、とりわけ合理的意思決定を重視する領域ではこの独自のリアリティ構成を行い、自らのモデルに従うエージェントという視点は近年まで比較的希薄であった。また同じ社会科学であっても主体のリアリティの構成を重視する領域、例えば現象学的社会学やエスノメソドロジーと社会システム論では、互いの概念に相互翻訳可能性がない状況が続いてきた。しかし近年、状況は大きく変化しつつある。様々なレベルの問題解決にそれに関与するエージェントが持つシステム理解そのものの調整や学習が必要であることが次第に認識されるようになってきたからである。複数のリアリティを調整し、相互主観的なリアリティを構成するようなシステムのデザインが現実必要とされている。この認識をモデルに転化することに関して、ハードなサイエンスの担い手を自負している人達には、恐らく不安があることだろう。

では実際に、複数のリアリティを扱うようなシステムの分析はどのようにして可能となるのだろうか。この問題についての洞察を得るために、主体を含む複雑なシステムから適当なシステムの性質を抽出してモデル化していく作業がどうしても必要となる。実際、複雑系に関する様々な研究の歴史は、ニュートン力学のパラダイムからは導かれない、複雑なシステムの様々な特性を定式化してきた歴史でもあった。

一例を挙げよう。今日我々は「目的」という概念を

フィードバック制御に関して何の違和感も感じずに用いる。しかし目的という概念は、本来変分原理による定式化という例外を除けば、ニュートン力学のパラダイムにはなじまない怪しげな概念であるとも言える。それでも生物や社会を扱う人々にとっては、決して無視することのできない概念であった。しかしいくら目的という概念がアリストテレス以外の目的因という古い歴史を持つとしても、それが科学的モデルを持つシステムの性質として確立されるには、フィードバック概念の定式化が必要だったのである。一度定式化が与えられれば目的という概念は、例え本来の含意の一部が定式化されたものであっても、安心して使うことが可能となる。

逆に言えば、数理としてのハードコアを持たないシステムの性質はどうしてもそのリサーチプログラムの展開可能性は限定されたものになる。これについては歴史的な事例がある。社会科学の領域では、ポジティブフィードバックの概念を用いて、今日自己組織化やシナジェティクスの名で一般的となった力学系の分岐理論による構造変動のパラダイムとほぼ対応するような議論を1960年代から行っていた。これは構造や構造変動という概念が、社会科学が扱うべき最も重要なシステムの性質の1つだからである。だが結局構造変動という概念に対する数理的ハードコアが得られたのは物理学の領域からであった。

同様に、主体によるリアリティの構成や、構成されたリアリティ（モデル）に依拠したエージェントの活動などの諸概念も、適当なモデルが与えられることで安心して利用可能な概念となるだろう。複数のリアリティを扱う科学を構成するにはそれに関するシステムの性質の定式化とその分析が不可欠となる。このようなシステムの性質の候補に内部モデルという概念がある。この内部モデル概念は、やはり長い歴史を持った概念であるが、まだ十分に定式化されているとは言い難い状態にある概念である。これについては、例えば木嶋がハイパーゲームを拡張する形でこの内部モデルの相互調整に関するモデルを扱っている。

我々はこの内部モデルを持つエージェントを扱う理論が、主体的なエージェントからなるシステムの分析では不可欠であると考えている。そのためには内部モデルに関連したシステムの性質の数理的分析が不可欠である。しかも内部モデルというシステムの性質は、自然科学的な対象の中ではまず見いだすことのないシステムの性質である。それゆえ社会科学的な立場から

のモデル作りが不可避となる。このような主体を含む複雑系のシステム分析というアプローチに対して、いわゆる Complexity に関する議論はどのように関連するのであろうか。

3. 複雑系という奇妙な流行

近年、複雑系という奇妙な流行が日本の思想界に始まろうとしている。複雑系という流行は、構造主義やサイバネティクス、大規模システム、ホロン、カタストロフィー理論、シナジェティクスや散逸構造といった自己組織化、カオス、オートポイエシス、人工生命などの流行と同様に何度目かの思想・学問的流行であり、それ自体はとりたてて言うこともない話である。ただし、複雑系という概念は、長い歴史を持つ普通名詞とでもいべき一般的な概念である。これが特定の実体のあるリサーチプログラムであるかのごとく喧伝され、そしてお定まりの流行の終焉のレッテルが貼られてしまうのであろうか。我々のポリエージェントというリサーチプログラムは、多主体複雑系の科学として構想しこれから長く研究を進めていきたいと考えている。その最中に突然生じたこの奇妙な複雑系のブームは、極めて日本的なものであるとはいえ若干のコメントが必要だろう。

Complex System あるいは Complexity をめぐる研究は、ニュートンのパラダイムにのらない残余項を定式化しようとする長い戦いの歴史である。複雑なシステムに関するパラダイムを求める動きは、10年ごとく新しい流行が生じる形ですっと続いてきた。本来複雑系あるいは complexity に関する固有のリサーチプログラムというものがあるわけではない。ニュートン力学的な時間発展方程式に関する問題関心の範囲を越えたシステムの性質をどのように定式化するかで複雑性に関する様々なリサーチプログラムが提起され発展してきたのである。

サイバネティクス、大規模システム論、エルゴード性、生態系の大規模システム解析、ゲーム理論、ポピュレーションダイナミクスによる生態系分析、カタストロフィー、自己組織化やシナジェティクス（力学系の分岐理論）、統計力学の拡散と発展、カオス、ESS と進化ゲーム理論、一般システム論、階層システム論、ライフゲームやリンデンバウムシステム、人工知能や分散人工知能、オートポイエシス、認知科学や認識人類学、セルオートマタの再流行、遺伝アルゴリズム、人工生命など1950年代くらいから現在まで、様々な領

域で何らかの形で複雑なシステムの性質を読み解こうと試みてきたアプローチやそのキャッチフレーズは多岐にわたる。

これらのパラダイムはまた相互に様々に影響を与えている。領域固有の問題意識をクロスさせながら、ニュートン的な力学系の時間発展の世界像からは出てこない様々な問題意識を人々は Complex Systems という一般的な言い方で捉え、それを領域固有のリサーチプログラムとして展開していった。その意味で Complex Systems 論という特定のリサーチプログラムがあるわけではない。

近年喧伝されることの多い、サンタフェ研究所の Complex Adaptive System というリサーチプログラムも進化に関する遺伝アルゴリズムとセルオートマタの理論をハードコアとし、方法論としてコンピュータシミュレーションを多用した1つの研究プログラムに過ぎない。社会科学の領域では、組織などのシステムがどのように外界の Complexity へ対処するかという問題設定を始めとして、主体を含む複雑なシステムへの問題意識はずっと継続されてきたものであり、何ら目新しいものではない。

これらの複雑なシステムの性質を抽出し定式化しようというアプローチの中で、もっとも大きな影響を与えたのが、恒常性あるいは限定された意味での目的という性質に定式化を与えた、サイバネティクスのフィードバック概念と、構造とその変動という概念を定式化した力学系の分岐理論に基づいた自己組織化やシナジェティクス概念であろう。これらはいずれも自然科学的な領域を対象としつつ社会科学の領域まで大きな影響を与えている。さらに、進化という概念を定式化しようとした遺伝アルゴリズムや進化ゲーム理論の進化安定 (ESS) という概念も大きな影響を与えつつある。だが、これら従来までの複雑なシステムに対するシステムの性質は、我々が着目する幾つかのシステムの性質の分析には不十分である。例えば主体的なエージェントは自身や他のエージェントおよび外界に対するモデル (内部モデル) をエージェント自体が持つシステムである。またエージェントは、交互に互いの活動や内部モデル自体を参照することで様々な活動の調整を行う。このような自律的なエージェントでは、その状態を直接に自由に制御する制御概念を適用するには限界がある。エージェントは自身の活動プログラムや目的に従って活動しており、その活動の境界条件を様々に変えてやることでエージェントは自身でその活

動を変更し結果として間接的にある状態が実現される (間接制御)。さらにエージェントは複数が機能的に連結し、さらに高次のエージェントを形成する (高次構造)。これら内部モデルやその相互参照、間接制御、高次構造といった諸概念を定式化する必要がある。だがこれらは分岐理論のように何らかの形で従来の物理的システムを規範としたアプローチではなかなか定式化することの難しい領域でもある。

我々がポリエージェントシステムという比較的ゆるいリサーチパラダイム、あるいはリサーチイニシアチブとでも言うべきものを提唱する必要があると考えているのは、この主体を含む複雑なシステムに関する理論の向かうべき方向が、これら自然科学的なパラダイムや研究プログラムの延長上にはないところにあるように思えるからである。

4. エージェント指向の複雑系分析

我々は自律的主体を含む複雑なシステムに対する諸アプローチがしだいにエージェント指向の科学に収束するであろうと考えている。すでにエージェント指向のアプローチとしてゲーム理論や、分散人工知能 (DAI)、ロボティクス、自律分散系、マルチエージェントシステムなど様々なアプローチがある。これらはエージェント指向といってもかなりの幅があり、現在のところそれらの間に用語や概念間の翻訳可能性が常に成立しているわけではない。また自律的エージェントに関する個々のリサーチプログラムの問題意識や従来からの複雑性に関する諸システムの性質との関係が擦り合わされているわけでもない。それゆえによけい、このエージェント指向のアプローチについて、そのリサーチイニシアチブとしての概念構成を行うことが重要となる。

このことを少し数理的に見てみよう。すでに述べたように、力学系の分岐という概念は、構造と構造変動という概念を定式化したという意味で画期的なものであった。しかしこのアプローチは、微分方程式レベルのものである。これがエージェント指向の活動モデルと結びつくためにはいまひとつ工夫が必要である。

それを成し遂げたのが、メイナードスミスの進化安定の概念であった。彼の進化安定の概念は、ゲームから導かれるある戦略を取る集団の人口分布に関する力学系 (レプリケータダイナミクス) の安定平衡解をゲームレベルで特徴づけることに成功した。エージェント間の相互作用の基本モデルとしてのゲーム理論と力

学系の平衡点の安定性解析を結びつけたのである。この分析は、ゲームというエージェント間の2主体間関係に基礎を置いて力学系を導出するという点で社会科学にとっても非常に重要なブレイクスルーと見ることができる。しかし奇妙なことにこのエージェントベースの力学系については、生物学や進化経済学の領域で平衡点の複数性や安定性が論じられることは多いが、構造変動のパラダイムとしての力学系の分岐理論とは積極的に結びつけられていない。そこで簡単な例をとってこのリサーチプログラムの拡張の可能性を検討し、その上でこの拡張でも扱えないポリエージェントの問題意識の意義について考えたい。

ここではメイナードスミスのものよりさらに簡単な、エージェントの態度変容モデルを考える。ここで扱うのはエージェント間の直接的な相互作用ではなく、外的環境条件（外場）の下でのエージェントの反応に基づく力学系である。ここでエージェントは2つの代替案 {1, 2} から1つの態度を選択すると仮定する。この選択はその代替案に関してエージェントが得る利得に比例して意思決定が逐次的に行われると考える。意思決定を行うエージェントの人口を N とすると、 N 主体のうち、どれだけが代替案1を選択し、またどれだけが代替案2を選択するかのダイナミクスがここでいうレプリケータダイナミクスとなる。これは人口 N を一定とすれば、領域 $[0, 1]$ のダイナミクスに正規化でき、さらに代替案1を選択する率を P とすれば、代替案2を選択する率は $(1 - P)$ となり、1次元の力学系と見なされる。

ここで各々の代替案を選択することで得られる利得を W_1, W_2 とする。平均利得 W^- を、 $W^- = \sum P_i W_i$ と置く。するとこのレプリケータダイナミクスは、 $dP/dt = P(W_1 - W^-)/W^-$ となる。ここで平均項で割った部分の方程式も質的に同型なので $dP/dt = P(W_1 - W^-)$ を解析する。

ここでは情報ネットワーク上に情報を積極的に提供する態度とフリーライダーを決め込む態度の間の戦略的意思決定モデルという解釈を与えよう。代替案1を情報提供戦略、代替案2をフリーライダー戦略とする。このときネットワーク上のプレイヤーに対しては、ある確率で評価が行われ、情報提供者は評価利益 a が与えられ、フリーライダーには評価損 e が与えられると考えよう。この評価確率を β とする。また情報の利用者としてどちらも d という利用利益を得ること、情報の提供者は提供のためのコスト c を必要とすると仮定

する。すると情報提供とフリーライダーという戦略の各々の平均利得は W [情報提供] = $W_1 = \beta(a - c + d) + (1 - \beta)(d - c)$, W [フリーライダー] = $W_2 = \beta(d - e) + (1 - \beta)d$ となる。これから情報提供者の人口に占める比率のレプリケータダイナミクスを計算すると、 $dP/dt = p\{W[\text{提供}] - W^-\} = p(\beta(a + e) - c)(1 - p)$ となる。この力学系の平衡点、つまり $dP/dt = 0$ となる点は $p = 1$ または $p = 0$ となる点である。つまりフリーライダーと情報提供者のどちらも平衡解となる。その安定性を見ると、 $p = 0$ の付近では $dp/dt < 0$ であれば0へ引き戻されるので、解は安定となる。つまり $\beta < c/(a + e)$ のときに系は安定する。 $p = 1$ の付近では、 $dp/dt > 0$ であれば1へ引き戻されるので、解は安定となる。つまり $\beta > c/(a + e)$ のときに系は安定することになる。

これは1次元力学系で●を安定解、○を不安定解で示すと、

$$0 \quad \text{○} \text{-----} \text{●} \quad 1 \quad \text{if } \beta > c/(a + e)$$

$$0 \quad \text{●} \text{-----} \text{○} \quad 1 \quad \text{if } \beta < c/(a + e)$$

と図示される。これは、構造パラメータ群 (β, a, e, c) の変化によって力学系が分岐することを意味している。 $\beta = c/(a + e)$ の点が構造不安定点である。このモデルは、評価される可能性が高いとき、つまりネットワークの中での評判が立つ可能性が利益に対する参加コストの比率よりも高いならば人々は情報提供的な態度へ急速に分岐（構造変動）するというを示している。このモデルは簡単ではあるが、外的境界条件に対するエージェントのミクロの反応からマクロのレプリケータダイナミクスを導出し、その構造変動を与える事例を与える。メイナードスミスから始まるゲーム論的に導出されたレプリケータダイナミクスでは、このダイナミクスをエージェント間の相互作用を与えるゲームの利得表から導出してその安定性解析を、利得表レベルで進化安定という概念を用いて行っている。だが彼らの問題関心には上述のような分岐による構造変動という問題意識はあまり見当たらない。これは生物の場合、特に構造パラメータを動かして平衡解を分岐させるという事例が少ないことも関係するのだろう。経済学に問題を限局する限りでは、制度的多様性をナッシュ均衡の複数性で説明する枠組みが普及しつつある。ナッシュ均衡解が、レプリケータダイナミクスとしての力学系の平衡解と結びつくならば、平衡解の複数性が制度あるいは構造の多様性ということになる。これは安定平衡解のいずれかが実現されている構造で

あるとする、力学系の分岐パラダイムでの解釈と一応対応する。ただし、力学系の分岐パラダイムでは、構造変動へ認識関心が向けられていた。この立場からは実現された平衡解としての構造を背後で規定している構造パラメータを制度的構造と呼ぶほうが妥当である。

いずれにせよこのモデルの場合も実は、ゲーム理論よりさらに単純な効用最大化のモデルとして、力学系を書くまでもなくパラメータに依存した最適意思決定問題として解くことができる。つまり力学系的な構造変動のモデルは、エージェントの相互作用や反応レベルで定式化したり、分析することができる。このようにしてエージェントベースの相互作用や外的条件に対する反応モデルを、分岐理論型の複雑系の分析パラダイムに結びつけることができる。ここで例えばエージェントの位置情報を入れれば空間項の入った自己組織化モデルになる。

それではこのようなモデルの上でポリエージェント的な問題意識は定式化可能だろうか。先に述べた間接制御の概念は、このような枠組みの中で部分的に論じることができる。エージェントは自身で最適化意思決定を行っている。その活動を間接制御するためには構造変数を利用して制御するしかない。だがこの方法では上述の力学系の例では、 $[0,1]$ の状態空間の任意の値に持っていくことはできない。0または1に力学系の状態を持っていくことができるだけある。また上述の意思決定をゲーミングの形で人間集団にやらせてみると、エージェントの主観的な思い込みや誤解（内部モデル）が意思決定にかなりの影響を与えていることがわかる。この力学系をいじって主観性を導入する方法もないではないがやはり限定されたアプローチとなる。やはり主体を含む複雑系として問題を見たとき、このような角度からのエージェント指向のアプローチは、一定の有効性を示すと同時にその限界をも明らかにしていると言える。社会科学では、その理論フレームと現象への当てはめの間にギャップがあり、解釈に関するアートや信念の共有を必要とする場合が多々ある。社会科学で使われているモデルは、現象から一定の距離を持った説明の型としてのモデル（フレームモデル）であると言ってもよいだろう。それゆえモデルの表現力の限界を見定め、現象の持つ多様性を覆い隠すようなモデルの適用に十分注意を払う必要がある。

ポリエージェントの研究プログラムを策定しようと試みる中で、我々は様々な形で上述の新しいシステムの性質の定式化の試みを行っている。例えばエージェ

ント間の内部モデルの相互参照を、ハイパーゲームという枠組みの中でゲーム理論的な立場から論じることが木嶋によって行われている。我々は今後この内部モデルやその相互参照概念をより一般的に拡張する必要があると考えている。内部モデルをゲーム以外の形で表現することができ、内部モデルの相互参照や、間接制御の定式化などが可能となる、十分な表現の多様性を持ったフレームモデルを定式化してその上でリサーチプログラムを具体的な形で示すことが我々の大きな課題である。

5. おわりに

先に述べたような様々なエージェント指向のリサーチプログラムは、今日複雑系に関する奇妙な流行の影に隠れて日本では比較的紹介されることが少なかったり、違う分野で別々に紹介され相互のつながりが認識し難かったりする。だが、エージェント指向のアプローチは複雑なシステムに関する脈々と受け継がれている研究の流れである。例えばロボティクスや自律分散型の制御ではエージェント指向のアプローチは1つの流れを作っている。同様にDAIでもエージェント指向は鍵概念である。また社会学の領域でもエージェント指向のアプローチは常に底流で動いている。

ただ、構造変動に関する複雑系の理論が、力学系の分岐理論というハードコアを得て、研究プログラムとして大ブレイクを遂げ、あるいは遺伝アルゴリズムがそのアルゴリズムとしての定式化の平明さから非常に広範にわたる応用領域を見いだしたのに対して、エージェント指向のアプローチは現在のところ確たるハードコアとなる数理的定式化を見いだしているとは言えない。先に述べた内部モデルや間接制御のような概念は、ポリエージェントシステムのリサーチプログラムの重要なハードコアを構成するはずである。これについては、ウォンハムから始まる内部モデルを持つシステムに対する制御理論研究の歴史があるが、現在までに決定的なブレイクスルーが得られたわけではない。しかし、数理的ハードコアが得られないということは、そのリサーチプログラムに意味がないということにはならない。今日の分岐理論に基づく散逸構造やシナジエティクスといったパラダイムが登場する以前から、ポジティブフィードバック概念を使って、構造変動についてほぼ同様の議論が行われてきた歴史がある。この歴史から我々が学ぶべきことは、数理的なハードコアを得るための事例としては自然科学や工学での事例

の方が確かに堅固であり、我々がエージェント指向の新しいパラダイムに関するリサーチイニシアチブを提案するにあたって、これらの領域との密接な関係を保って研究を進めることが重要であるということであろう。意味や解釈を扱うソフトなアプローチと工学的なハードなアプローチは敵対するものではないのである。

文献

- 青木昌彦, 奥野正寛編著, 『経済システムの比較制度分析』, 東京大学出版会, 1996
- 出口 弘, 複雑系の認識原理と自己組織化, 東京工業大学総合理工学研究科システム科学専攻博士論文, 1986
- 出口弘, 自己組織化研究の方法論批判 科学革命としての自己組織化, 現代思想, VOL.16-1, 1988
- Hiroshi Deguchi, "MULTI AGENT ECONOMICS AND ITS GAMING SIMULATION", in 『Modeling and Control of National and Regional Economies 1995』, edited by Vlacic et al., pp.269-274, Pergamon Press, 1996
- 出口 弘, 情報と複雑系の経済システム論, 経済セミナー1月号, pp.32-39, 1997
- 出口 弘, ゲームとプログラミング社会科学, ゲームプログラミング bit7月号増刊, 1997
- D.E.Goldberg, 『Genetic Algorithms in search, Optimization, and learning』, Addison Wesley, 1989
- J.H.Holland, 『Hidden Order - How Adaptation Builds Complexity』, Addison-Wesley, 1995
- 木嶋恭一, 出口弘編著, 『システム知の探求 1 - 決定するシステム-』, 日科技連出版, 1997
- 木嶋恭一, 『交渉とアコモデーション』, 日科技連出版, 1996
- メイナード・スミス, J.著, 『進化とゲームの理論』, 寺本英, 梯正之訳, 産業図書
- 高木晴夫, 木嶋恭一, 出口 弘, 他著, 高木, 木嶋, 出口監修, シリーズ社会科学のフロンティア, 『マルチメディア時代の人間と社会』, 日科技連出版, 1995

新時代のコンピュータ総合誌

隔月刊

Computer Today

偶数月18日発売/本体905円

9月号・特集

複雑系とコンピュータ

——応用分野広がる非線形科学——

神経型システムの記憶情報処理	矢内 浩文
カオスニューロコンピュータ	堀尾 喜彦
ブレインウェア：連想記憶と神経競合	深井 朋樹
ニューラルネットワークを使った 手書き文字認識	猿田 和樹
遺伝的アルゴリズム	田中 雅博

連載 CMC研究ノート インターネットと法 続・アルゴリズムの工具箱 他

月刊誌

数理科学

毎月20日発売/本体952円

10月号・特集

「高温」超伝導理論の新展開

まえがき	福山 秀敏
強相関の立場から	永長 直人
電子状態について	前川 禎通
超伝導が起こるまでの異常金属相 フェルミ液体とスピンゆらぎ	三宅 和正
超伝導磁束状態の不思議	立木 昌・胡 曉

Java ライブラリ

●日本サン・マイクロシステムズ(株)編●

Java プログラミング

① インタフェース総説

——拡張APIとJDK 1.1——

B5・144頁・本体2000円

Java プログラミング

② JDBC

B5・208頁・本体2400円

[好評発売中]

サイエンス社

〒151 東京都渋谷区千駄ヶ谷1-3-25 ☎(03) 5474-8500

インターネットホームページ

<http://www.bekkoame.or.jp/~saiensu>

*表示価格は全て税抜きです。