計算能力の進化と集団の形成

神戸大・自然科学研究科　　橋本　敬[*]

1　計算能力と Chomsky の階層

「進化」という言葉は、無造作に使われていることが多いが、ある見方をしたとき、あるいは何らかの軸を設定したときに初めて「進化」の話ができるだろう。では、バクテリアから人間に「進化」したと言った場合、何をもって「進化」というのであろう。一つの見方として、たとえば「知能」という見方をしてみよう。人間中心的な考えではあるが、一般的にバクテリアよりは人間の方が「知能」が高いと考えられている。「知能」がどのようなもので計られるかは今はまだ分からないが、一つの側面として計算能力というものがある。人間の知能が単に計算能力によって計れるものではないことは当然である。しかし、計算能力という観点もあるのではないだろうか。計算能力を軸として、それが上昇するような進化のシミュレートを考えてみたい。

計算能力は計算理論の分野においてオートマトンの計算能力という形で研究が進められている。オートマトンの計算能力の階層は、形式言語の語の受理能力を分類した次のような Chomsky の階層によって特徴付けられる。

- (0 型) 句構造文法
- (1 型) 文脈依存文法
- (2 型) 文脈自由文法
- (3 型) 正則文法

ここで、上に書いたものほど言語の受理能力が高い、すなわち多くの言葉が解かるということである [1]。

同じ型に属する文法でも発話・理解できる語は異なる。様々な異なる文法を持ったエージェントがネットワークをつくる状況における会話とはどのようなものだろうか。たとえば、

- 他のエージェントがあまり話さない語を発話する
- 他のエージェントに理解されない語を発話する
- 他のエージェントの発話する語をより早く理解できる

という設定をしたとする。この場合、エージェントのつくるネットワークの間ではどんな語が話され、どのような文法が発展してくるだろうか。こういったことをコンピュータ上にモデルとして構築することにより考察してみたい。

---

[*] e-mail: toshiwo@gradis.scitec.kobe-u.ac.jp

2　会話ネットワークにおける進化モデル

2.1　書き換えルールシステムの会話ネットワーク

書き換えルール・システム（以下 RRS と記述）として、個々の文法を書き換えルールとして持っているようなエージェントを考える。書き換えルールは $X \rightarrow A$ の形をしており、Xはひとつの非終端記号であり、A は１つ以上の終端記号または非終端記号の列である [2]。個々の RRS はそれぞれの持つ文法を順次あてはめることにより語の発話を行ない、全ての RRS がルールを右辺から左辺へあてはめることにより発せられた語の理解を試みる。

このような RRS を 10 個用意し、順番に１個ずつ発話しその語をみんなで聞くという「会話ネットワーク」をつくる。

2.2　進化のダイナミクス

RRS の進化の指標として次のものを得点とする。

$$P_{tot} = R_{sp}P_{sp} + R_{rec}P_{rec} + R_{br}P_{br}$$

P_{sp}、P_{rec}、P_{br}はそれぞれ発話による得点、理解による得点、そして、聞き手（発話した RRS 自身も含めて）に理解されたときの得点である。また、R_{sp}、R_{rec}、R_{br}は各得点の結合比である。ここでは $R_{sp} = 3.0, R_{rec} = 1.0, R_{br} = -2.0$ とする。$R_{br} = -2.0$ なので、聞き手に理解されると点が下がる。よって、より多様な語を発話・理解する方向へと RRS が進化することが予想される。

各タイムステップで、ある割合で RRS の持つルールを変異させた新しい RRS をつくる。ここでは、付加（新たなルールをルールリストの最後に付加する）、置換（新たなルールと古いルールを入れ換える）、削除（ルールをひとつ取り除く）の３種類を考える。ルールの変異のさせかたも１文字の付加、置換、削除の３種類がある。このようにしてできた新しい RRS を得点の低かったものと入れ替え、全体で 10 人の RRS がいるようにする。

3　アルゴリズミックな進化と共語集団の形成

ランダムに作った最小のルールから出発した RRS 達が発展し、より高い発話の多様性を持った RRS が出現する（図１）。RRS の持つルールにより作られる語を生成木の形で描いてみよう（図２）。初めに存在する RRS は生成木が sequencial な形をしている (a)。すなわち、$S{\rightarrow}0$ のような単なる書き換えである。付加的突然変異により新たなルールを獲得し、別の語への書き換えが行なわれる。これにより、生成木は branch 構造になる (b)。さらに、branch のしたでさらに branch がある（multi branch）構造が現われる (c)。

初期の発展では新たな RRS が得る多様性はそれほど高いものではなく、たかだか数個の新たな語を発話・理解できるようになるだけである。しかし、図１をみると $t = 192$ のところで、急に高い発話の多様性を持ったものが現われている。このときの生成木の変化は図３のようになっている。変異前のものを見ると多くの branch がある。そこに $X{\rightarrow}00$ というルールが付け加わることで、たくさんの multi branch がつくられる。このとき $X{\rightarrow}00$ というルールは、他の多くのルールにくっついて多くの新たな語を生成するのに活躍する。これを生成木のモジュールと呼ぼう。このようにモジュールとなるルールを一つ得ることにより発話・理解できる語が飛躍的に増える。

さらに $t = 310$ で、発話できる語の多様性が高い RRS が登場している。モジュール型の進化までは、生成木が複雑になっていき葉が増えていく進化であったが、ここでの変化はこれまでのものとは違いルール

の中でループを形成するものができるのである。（図４）ループを作ると無限に書き換えが可能になり、生成木ではそのルールから作られる語をすべて表すことができない。

しかし、モジュールを持ち多様な語を発話・理解できるRRSが現われてもしばらくすると死んでしまう場合がある。たとえば、先ほど見た$t = 192$でモジュール的に進化したRRSは、その後、より発話の多様性の高いRRSが現われるわけではないが淘汰されてしまう。また$t = 310$以前にもループを持ったRRSは現われており、発話の多様性も高いものであったがすぐに淘汰されている。

特に顕著な場合が、$t = 403$のあたりで、系図のこのあたりを拡大したものが図5である。これを見ると発話の多様性の高いものがつぎつぎに淘汰されている。このように発話・理解の多様性が高いものが淘汰されるのは、共語集団(Common Word Ensemble)が形成されているからである。共語集団とは、特定の語を発話・理解しあう集団である。$t = 403$では五つの共語集団が形成されている。その中で小数派のRRSは、発話・理解の多様性が高くても淘汰される。この共語集団中では、モジュールやループといった個々の文法の持つ高い能力より共語集団の形成が進化的に有利になるのである。

4　共語集団 $\rightleftharpoons$ 二重分節

たくさんの要素が相互作用している系で、その要素の動きが何らかの秩序を持つという階層的な構造はどのようにして進化してくるのだろうか[3]。例えば、協力関係などの社会性の進化には弱いカオスが関係していたり[4]、ジレンマゲームにおけるTFT的な戦略が必要であると考えられている[5]。ここでは、モジュールを持つことや、ツリー構造よりもループ構造が多様な語を発話・理解できること、さらに共語集団ができることが進化の方向性を決めている。その中で「二重分節」という文法の秩序化を進化させる。

二重分節とは、アルファベットや五十音のような有限個の記号を用いる自然言語で、無限に多様な文をつくることを可能にしている仕組みのことである。すなわち記号を組み合わせて「語」をつくり、その語をさらに組み合わせて「文」をつくるという仕組みによりこれを可能にしている。ここでのRRSの持つ文法にも同じような構造ができてくるのである。共語集団が作られてからのRRSのもつ文法をみると、

1. 非終端記号 → 終端記号列
2. 非終端記号 → 非終端記号列、
   非終端記号 → 終端記号と非終端記号の組み合わせ

の2つの形のルールがある。受理過程ではこの矢印を逆向きにして、右辺から左辺へと書き換えを行なっていき初期記号Sまで書き換えができれば理解したことになる。1の形をしたルールは共語集団内でよく話される語を作っている場合が多い。共語集団においてよく話される語はより早く理解できたほうが有利なので、そのような語を受理する過程ができるだけ短くなるような文法になる。つまりよく話される語は、1の形のルールの特別な場合である

$$1'.\ \ S \rightarrow \text{よく話される語}$$

というかたちのルールによって理解する。あまり話されない長い語は、2の形のルールを用いて1の形のルールで作られた語を組み合わすことにより理解しようとする。これは自然言語の「二重分節」に対応するとみることができる。すなわち前者の形のルールは「単語」をコードし、後者のルールは「単語」を組み合わせて「文」を作っていると見ることができる。RRS間に上位の構造が造られることにより、それを反映した形でRRS自身の持つ文法も秩序化してくるのである。さらに、文法は共語集団に合わせて秩序化するので、その共語集団を安定に保つ効果を持つ（図６）[6]。一方が生産者でもう一方が生産物であると

いう一次元的な図式ではなく、自律形成においては生産物がまた生産者を生産するという循環的な図式になる。

しかし共語集団内では発せられる語が特殊化し、多様な語が話されにくくなり RRS の能力は低くなってしまうかも知れない。つまり全ての語を 1' の形で持てばもっとも早く理解できるだろう。これは Chomsky の階層では正則文法になる。実際はそうならずに少ないルール数でより多くの語を発話・受理するために、モジュール構造、ループ構造を造る。これは一方で、より受理能力が高くなるようにドライビングフォースがかけられているので、共語集団の中からさらに能力が高い RRS が生まれるからである。ループ構造は Chomsky の階層では文脈自由文法である。共語集団という上位の構造の創発と高い論理階層への進化とが、「二重分節」により実現されているのである。

ここで提出したモデルは、「言語」をモデル化したものではない。しかし、同じことばをやりとりする集団ができるという、人間社会におけるコミュニティーのでき方とよく似た振る舞いが見られる。このような違う論理で組み上げた人工システムから何が言えるだろうか。そこで、言語との対応について少し考えてみよう。言語には二つの側面がある。一つはコミュニケーションの道具としての言語である。これを外言語と呼ぼう。それにたいして、外界をうつしとり自分の内部に虚構世界をつくりだす能力が言語にはある。これは内言語と呼ぶ。外言語として発せられた言葉は発話者の内部世界から出たものであるが、それによりまた発話者自体の内部世界を解体・再構築する。そして、その言葉を聞いた者もこれに対して内部世界をつくるのである。すなわち、外言語として出された言葉も内言語としての働きを持つのである。つまり外界をうつしとり内部世界をつくることを絶え間なく行なうのだが、その過程で外界の事物に対して記号を割り当てる。そして単に記号を割り当てていくだけではなく、さらにそれらの記号間の関係を構築していくのである。

外界から虚構世界をつくるやり方を「文法」と呼び [7]、これを形式化したものが形式文法システムであるというふうにとらえた場合、記号の割り当てと記号間の関係の構築は、共語集団ができるにつれて、共語集団内でよく話される語をコードするルールと文をコードするルールに文法が構造化してくることに対応する。これは言語の習得過程、あるいは言語の発生過程で対象に記号を割り振っていくことから、それらの間の関係を作るという上位レベルでとらえられるようになることと関係し、Chomsky の階層の上昇はメタレベルでの理解ができるようになることにあたるであろう。また、同じことばをやり取りして集団を作っているということは、コミュニケーションが成立しているように見えるが、その共語集団のなかのエージェントの持つ文法を見ると、それぞれ違った文法を持っており、常に変化している。ことばがとびかうだけで、相手がどのような「文法」を持っているかは知りえず、どのような過程によりその語を受理しているかは分からない。つまり、お互いに話をしていても相手がどのようにしてそれを内在化しているかは知りえないのである。我々のコミュニケーションにおいて、全く知らない言語や外国語を話す人とコミュニケーションをとる状況で似たようなことを感じることはあるだろう。しかし、同じ言語を話しているもの同志でも実は状況は変わらないのではないだろうか。ことばがとびかうだけで相手がどのように理解しているかは知りえないという、本質的にディスコミュニケーションな状況なのである。

参考文献

[1] Hopcroft, J. E. , and Ullman, J. D. *Introduction to Automata Theory, Languages and Computation.* (Addison-Wesley Publishing Co. Inc. Reading, Massachusetts, U.S.A.(1979))
（野崎昭弘、高橋正子、町田元、山崎秀記訳：オートマトン　言語理論　計算論　I,II , サイエンス社 (1986) ）
語を受理するとは、ある文法によりその語（記号列）が導出されるということである。この階層において３型は２型に、２型は１型に、１型は０型に制限を加えたものだから、番号が小さい型の文法ほ

ど導出される記号列の集合が大きい。

[2] 書き換えルールの左辺を非終端記号１文字に制限することはRRSのもてる文法の型を正則文法、文脈自由文法に制限し文脈依存文法、句構造文法は現われないように設定したことになる。これは主に構文解析に関する技術的な問題による。

[3] Kauffman, S. A. "Requirements for evolvability in complex systems" in *Emergent Computation* (The MIT Press, Massachusetts U.S.A. (1987))
Kauffman, S. A. , Sci. Ame. , Aug. (1991) 78

[4] Kaneko, K. and Ikegami, T. , Phisica **D 56** (1992) 406

[5] Maynard, S. J. *Evolution and Theory of Games* (Cambridge University press, Cambridge, U. K. (1982))
（寺本英、梯正之訳：進化とゲーム理論、産業図書 (1985)）

[6] RRSはそれぞれの特長を持った語を発するが、特定の語を理解できない。共語集団が一つの種を構成すると考えればこれはアイゲンの提出する「準種」の概念に相当する。一つの種内に様々なアルゴリズムを用意し種内である程度のdiversityを保ち共語集団内の個々のRRSの持つ欠点を補いあい他の種からの攻撃に対処したり、種としての特定の同一性をもちながら様々な方向への進化の道を確保するのに役立っている。生物の進化では、ある方向への進化と多様性を広げるような進化が拮抗している。「準種」に関してはM. Eigen (1993) SCIENTIFIC AMERICAN, July, 日経サイエンス1993年９月参照

[7] この意味での「文法」はチョムスキーが言う意味での「文法」とだいたい同じである。しかしチョムスキーはその「文法」は誰もが生得的に持っており、普遍的であり、変化しないとした。ここでは「文法」は本来的に変化するものであり、個々人により違うもので、そのでき方は社会により大きな影響を受けると考えている。Chomsky, N. *Reflections on Language* (Pantheon Books, New york, U.S.A. (1975))（井上和子、神尾昭雄、西山佑司訳：言語論、大修館書店）

図１ :横軸に時間、縦軸に発話の多様性、すなわち全ての語のうちRRSが発話できる語の割合をとったもので、初期に発話できるRRSを「始祖」としそこから別れて出ている線はその変異体なので「系図」と見ることができる。線が途中で切れているものはその時点で淘汰されたことを示す。初期は非常に低い多様性であるが所々で高い多様性を持ったRRSが現われている。しかしその時点でもっとも高い多様性を持つにもかかわらず、淘汰されているRRSがいることも分かる。

図２ :RRSの初期の発展の様子を、文法を生成木の形で書くことにより示したもの。付加的な変異によりRRSの発話できる語が増える。その過程での生成木の形の変化を見ると、a) 順次構造、b) 分岐構造、c) 多重分岐構造というように発展している。

図３ :モジュール的進化の例。一つのルールがたくさんのルールから使われ、新たな語を生成する。このとき新たなルールを得る前の状態は非常に多くの枝を持っている。新たなルールを一つ得ることにより枝の先でさらに分岐し、発話・理解できる語の数が飛躍的に増える。

図４ :ループ構造の出現の模式図。$*X*$は右辺にXが含まれていることを表している。これに、$X \to 1Y0$というルールが付け加わると$*X*$と$*Y*$のあいだでループを形成し、無限に書き換えが可能になる。このように、無限に書き換えが行なわれる場合、生成木ではそのルールから作られる語をすべて表すことができない。

図５ :$t = 390 \sim 435$での系図。$t = 403$あたりで発話の多様性が高いものが淘汰されている。これは共語集団の効果によるものである。そして$t = 415$あたり以降、共語集団の中から多様性の高いものが現われてくる。

図６ :いくつかのRRSが共語集団をつくり、それにより文法が共語集団内でよく話される語を作るルールと、文を作るルールに構造化する。その文法の構造化により共語集団が安定になる。共語集団の形成と文法の構造化は循環的な関係にある。