

Title	マルチエージェント・モデルによる共通文法の獲得
Author(s)	中村, 誠
Citation	
Issue Date	1997-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/1041
Rights	
Description	Supervisor: 東条 敏, 情報科学研究科, 修士

マルチエージェント・モデルによる共通文法の獲得

中村 誠

北陸先端科学技術大学院大学 情報科学研究科

1997年2月14日

キーワード: multi-agent model, language acquisition, LTAG, Genetic Programming.

自然言語の文法の形式化は N.Chomsky 以来盛んに行なわれてきた. 自然言語文を処理するためには効率の良い形式化が求められるが, それをすることにより, 自然言語のもつ本質的な特性は極力損なわれてはならない. それは, 自然言語の文法は形式言語のそれとは違い, 統計現象であるということである. なぜならば, 未知語や未知の文法に遭遇した際に柔軟に対応して処理しなければならず, それは自然言語処理をするうえで, このような現象は頻発するからである.

いかに柔軟な文法表現をするかという問題について, 本研究ではどのようにして統計的な自然言語文法が発生するかという点に着目した. そのためには現在話されている自然言語の文法が, いかにして出現したかというその歴史を考察しなければならない.

なぜ, ある言語では語順が固定しているが, これが全ての言語には起きないのか. また, それが言語の歴史のある時期には見られても, なぜ, すべての時期には起こらないのか. これらの疑問を解く鍵の一部として, 近年ピジン, クレオールといった現象に対しての研究が盛んに行なわれている.

ピジンとクレオールの特徴を簡単に述べると以下ようになる.

pidgin 共通の言語をもっていないが, 通称その他の目的で互いに話をしたいと思っている人びとの間に発達した伝達のシステムである. pidgin は, そのもとになった言語に比べて, 語彙は限られており, 文法構造は単純化され, 機能する範囲ははるかに狭い. pidgin を母語とする者は誰もいない.

creole creole は, ある共同体の母語となった pidgin のことをいう. この定義は, pidgin と creole が言語発達という単一の過程における2つの段階であるという点を強調したものである.

共通言語を持たない2つ以上の集団が、一定期間以上、しばしば特定の目的を持って交渉を持つ際、何らかの接触言語が形成される。ピジンとは接触言語のひとつのタイプであり、いずれの集団の母語でもなく、文法的・語彙的な簡略化が著しく、使用する場面、用途が限定されている。

ピジンは、本来それを必要とする状況がなくなれば、自然消滅することが多い。しかし、場合によっては、異なる社会状況のもとで存続し、語彙的・文法的に拡張され、安定的に使用される地域共通語 (stable pidgin, expanded pidgin, lingua franca) となったり、また、それを母語として話す人々の登場によりクレオールとなったりする。ピジンの中には、異言語間の伝達的手段として非常に有用になり、日常的な補助言語として、それまでよりも公に認められた役割を果たすようになったものもある。さらに、共通語としての公的な地位を社会によって与えられることもある。これは、「ピジンの拡張」といわれ、それは、使用者の要求に答えるために新しいかたちが付け加えられ、以前よりもはるかに広い範囲の状況で用いられるようになるからである。

このような頻繁に起こる変化によって、先に述べた形式言語では処理しきれない自然言語のもつ不自然さが現れたのである。

以上のことを踏まえて、実際に起こった自然言語文法の淘汰する様子、すなわち互いに異なる文法を持つコミュニティが、コミュニケーションをとることを目的として、互いの文法を獲得する過程をシミュレートすることができるならば、自己改編するような文法表現をすることが可能ではないかとここでは提案する。

本研究の位置づけとしては、自らメッセージを発信できる、能動的で、自律的なエージェント間の協調、競合または組織化だと考えられる。そして、本研究の目的は、ある文法を持ったエージェント群の中に、異なった文法を持つエージェント群が加わったときに、多数のメッセージ交換を通じて、試行錯誤しながら、各エージェントが群の中での共通な文法を獲得していく過程を実現することである。具体的には、世界中に存在する pidgin, creole といった混合言語が発生する過程を、シミュレーションモデルとして実現することである。

エージェントが持つ文法には CFG ではなく、LTAG を用いた。本研究では LTAG に深層格を要素として定義した概念と組み合わせた LTAG を提案した。また、学習機構には遺伝的プログラミングを採用し、LTAG の持つ構造的特徴に対応するような学習機構を持たすことができた。これらの要素を組み合わせることにより、共通文法獲得モデルを提案し、並列環境においてその実験を行なった。