

論文の内容の要旨

論文題目

Numerical Studies of Language and Discourse Complexity:
A Dynamical Systems Approach
(言語と談話の複雑さに関する数理的研究 —
力学系によるアプローチ)

氏名 猪狩 一郎

本研究では、言語によるコミュニケーションを、二体のエージェント間の談話として、コンピュータシミュレーションにより数理的に扱う。

自然言語を伴う現象が数限りなくある中で、我々が談話に着目するのは、これが単文の解析に帰着しきれない現象だからである。一般に、自然言語のうち数理的解析に相性が良いのは、生成文法でいう言語能力の領域であるとされている。本研究は、言語能力と言語運用の両方にまたがっている談話という現象を、構成論的にモデル化し、この抽象モデルを通じてその数理的構造を見出そうというものである。

ここで構成論的というとき、言語でコミュニケーションする一般的なマルチ・エージェントモデルに対する批判的な主張が含まれている。これら言語によるコミュニケーションに利得表が与えられているようなモデル(ゲーム論的枠組)においては、言語システムとエージェントが自己組織的に共進化する、というシナリオを描くことが多いが、進化の方向はまず例外なく、系全体を複雑化するドライビング・フォース、すなわち利得表の作り方に暗黙のうちに埋め込まれている。また、このようなモデルの多くで、言語はエージェントの中に在る、「あること」を表現するための単なる伝達プロトコルとして形式化されている。このような枠組においては、「伝達事項がうまく互いに理解できること」に高い得点を与えて進化させれば、なるべく同一のプロトコルをもつ、均一な集団ができあがってそれが安定な固定点となるのは明らかである。そこで我々は、談話プロセスに「プロトコルのすりあわせ」以外の意味があることを表現するために、第三者的な談話の成功/不成功を得点付けしないで発話し合う、抽象的な数理モデルを提案する。

モデル化にあたっては、各エージェントはダイナミカル・レコグナイザーと呼ばれる、シンプルなりカレント・ニューラルネットワークによって構成する。ダイナミカル・レコグナイザー単体では時系列入力の中の構造を学習でき、有限オートマトンを模倣するような学習が可能であることが既に知られている。

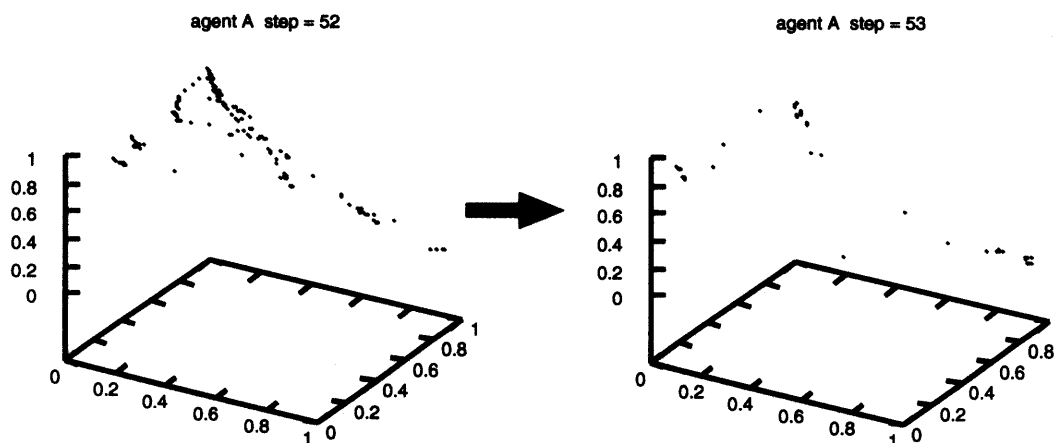


Figure 1: コンテキストプロット。雲状から島状にネットワーク構造が変化している

このエージェント (ダイナミカル・レコグナイザー) を独立に二体用意し、談話プロセスを無限ループさせる。エージェント A にとっての一回の談話ステップは、次のようなプロセスから成る。

1. エージェント A が (自分の発話と相手の発話を結合させた形で) 過去の発話列の一部を学習する。これは談話を自分なりにモデル化するということである
2. 次の発話用に、上のモデルを使って、“smoothness principle” になるべくよく従うような発話を選ぶため、全ての発話候補について計算を行う
3. 発話する

以上のプロセスをエージェント A、B が同時に行い、また始めに戻って無限ループするものとする。

我々のモデルでは、「何を発話すればよいのか」ということに関して、エージェントの上のレベルにあって客観的な評価をするメカニズムが組み込まれていない。そのかわり、上記プロセス 2 にある “smoothness principle” なる原理を天下り的に導入する。この原理に従って発話候補の「望ましい」度合を計算するのだが、その計算は「自分だけの談話のモデル」のみに基づいている点が大変重要な所である。いわば、エージェントは自分の思い込み (自前のモデル) に従って次の発話を選ぶのである。

発話候補は、簡単化のため 6 種類のアルファベットから成る (baa)(dii)(guu) のようなユニットチャンク 4 個の組合せの中から選ぶという制限を設け、2 種類の smoothness principle について談話のダイナミクスを数値シミュレーションにより分析した。

特に、「現在学習されているモデルに最も適合する発話を選ぶ」という smoothness principle を適用したときについては、発話時にランダムノイズが付随する場合も併せて調べた。このとき、発話パターンの安定性と、安定相から不安定相への移りかわりの機構を、ネットワークのコンテキストノードの状態と、発話候補のエラー構造とから特徴づけることができる。