

自然言語とプログラミング言語から見る言語の特徴

平野 正徳

東京大学 教養学部 (前期課程) 理科一類 1 年

I. はじめに

近年、グローバル化が進み、ますます他国との交流が進み、母語を異とする人同士の交流が活発となり、母語以外の言語を利用する機会も多くなった。しかしながら、言語の使用者数に大きな変動は見られないと考えられる。一方で、プログラミング言語においては使用数の順位変動が頻繁に起きている。この原因はプログラミング言語が主にインターネットを媒介として広まっていくのに対し、自然言語は人と人の直接のコミュニケーションで広まっていくという点で土地拘束が発生することと、プログラミング言語は自然言語に比べて、各言語間での共通点が多く、似ている点にあると考えられるので、これらの条件だけを変化させることで自然言語とプログラミング言語のユーザー数の変動を再現できるのではないかと考えた。この研究を通して、共通語に適した性質や言語の使用を支配する要素を分析することで、さらなるグローバルなコミュニケーションの手助けとなるかもしれない。

II. モデルの作成

前節で述べた通り、土地拘束の有無、言語間での共通点をパラメーター化した親和度を含めて考慮に入れ、*artisoc* のモデルを作成した。登場するエージェントは「話者」及び、言語の発信源となる「terminal」であり、空間は二次空間であり、様々なパラメーターを設定可能となっている。集計項目は各言語の各ステップでの使用者数と習得者数とした。

III. パラメーターの検討

ここまでのモデルで使用しているパラメーターの設定には検討が必要である。そこで、実際のユーザー数の変化に似たようなものになるようなパラメーター値を探すことにした。この際には、ブレイクアウトしないことや、首位言語の使用者数順位低下に関することを条件として絞っていった。

A. 自然言語の結果

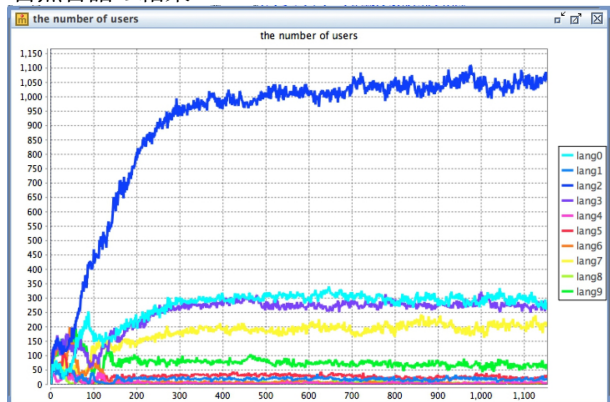


Fig. 1. 周りを見て言語を選ぶ割合:0.4 の時の自然言語のユーザー数のシミュレーション結果 (試行回数 1 回目)

各条件で 5 回ずつ試行したところ、おおよそ周りを見て言語を選ぶ割合は 0.6 以下とわかった。ただ、これでは周りを見て言語を選ぶ割合の値が分からないので、首位言語のシェアでフィッティングを行うことにしたが、明確な相関を見出せなかったため、この方法ではできなかった。

B. プログラム言語の結果

周りを見て言語を選ぶ割合が 1.0 でない場合はどれも言語間の親和度によらず、ブレイクアウトを起こした。

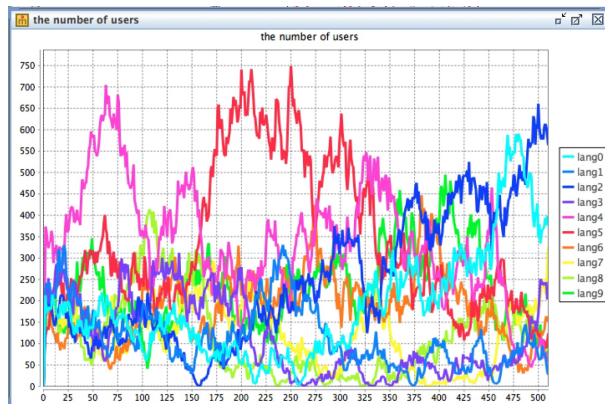


Fig. 2. 周りを見て言語を選ぶ割合:1.0 の時のプログラミング言語のユーザー数のシミュレーション結果 (試行回数 1 回目)

IV. 考察

自然言語における周りを見て言語を選ぶ割合は具体的に決定できなかったものの、おおよそ 0.6 以下である一方、プログラミング言語においてはおよそ 1 だとわかった。また、プログラミング言語において、親和度は直接的に影響はなかった。つまり、自然言語とプログラミング言語における違いは土地拘束の存在以外には周りを見て言語を選ぶ割合しかないということになる。それぞれの言語の親和度が高くなると、人々が使用言語を変更することがたやすくなると考えればつじつまが合う。

また、プログラミング言語において、周りを見て言語を決める確率が低い場合の結果に基づくと、自然言語は人々の移動距離が大きくなるとブレイクアウトしていき、言語統一が起こるかもしれないが、居住地という概念があるうちは極めて厳しいと考えられる。オンラインなどの仮想空間が広がれば統一言語ができるかもしれない。

今後の研究としては、自然言語における周りを見て言語を選ぶ割合の検討が考えられる。今回は首位言語のシェアで検討しようとしたが、周りを見て言語を選ぶ割合を変化させて、明らかに現れる変化は言語が空間全体に浸透するまでの step 数の増加である。これをうまく使えば周りを見て言語を選ぶ割合を決めることができるかもしれない。

また、プログラミング言語は今回、完全に土地拘束のないものとしたが、実際にはインターネット上にもコミュニティーが存在するわけで、完全にないとはみなすことはできない。このような要素が抜けてしまっているためにプログラミング言語における周りを見て言語を選ぶ割合が 1 になってしまっている可能性が高い。ただ、これは土地拘束が完全にあるとみなした自然言語において、首位言語の陥落が頻繁に観測されたのは周りを見て言語を選ぶ割合が 0.9 以上の場合であったことを鑑みれば、プログラミング言語の周りを見て言語を選ぶ割合は極めて高いと考えることは間違いないと思われる。