

第2回: チューリングとチューリングマシン

チューリングマシンで有名なアラン・チューリングを紹介する。チューリング・マシンは、現在のコンピュータの基本的なモデルであるとともに、言語処理の基本的な構造も提供している。関連して、マルコフ連鎖 (Markov Chain) についても紹介する。

1 アラン・チューリング

アラン・チューリング (Alan Mathison Turing) はイギリスの数学者で現代計算機科学の父と呼ばれる。チューリング・テストは、人工意識 (機械が意識を持ち、思考することができるか) の議論に挑発的かつ大きな影響を与えた。チューリングは、理論的に証明できることは手続き的に書き下すことができると考えた (有本 2000: 25)。



Alan Turing 1912-1954

2 チューリングと言語学

チューリングは言語学者ではない。しかし、チューリングの思想は言語学に有益である。直接的にはあまり言われないことであるが、チューリング・マシンとマルコフ・チェインは、隣接する記号の読み取りという点では同じである。この2つの形式化はリニアな言語音声人間が読み取る点において酷似している。

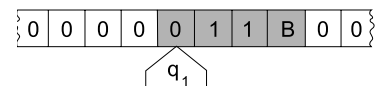


図 1: チューリングマシンの原理図

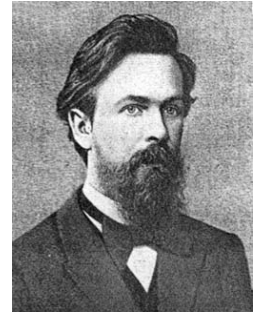
チューリングマシン

アラン・チューリングが 1943 年に、今日のコンピュータの基礎概念といわれるチューリングマシンという概念を英国で発表しました。図 1 のようなもので、テープの上に書き込み読み出しのできるヘッドを置き、テープの上では一駒ごとに情報があり、機械側に取り込んだり、書き込んだりして、機械の内部状態を変え、その上で、ヘッドをテープ上で一駒だけ動かすことができるという簡単なものです。この装置で、テープに無限の長さがあれば、コンピュータが行うような演算の総ては可能であるというものです。そういわれれば、そんな気がしますが、全くの原理としての先見性であって、実際に作られたわけではありません。コンピュータが実際に作られたのは、それから約 10 年後です。磁気記録媒体としても、ワイヤ型の録音機は 19 世紀からあったそうですが、磁気テープによる録音機は、その 2 年後 1936 年にドイツで発明されましたが、連合国側が知ったのは戦後のことだそうです。チューリングのこのマシンによって、アルゴリズムという概念が確立されたのだそうです。現在のコンピュータも突き詰めれば、このチューリングのマシンの原理に従っているといえるというのが定説です。 「コンピュータと脳」深町一彦より^{*1}

- 問 1 日常生活で図に似た構造のものを探し、それとチューリングマシンはどんな点で似ているだろうか。
- 問 2 言語をチューリングマシンでモデル化するならば、「ヘッド」「テープ」「セル」はどんな言語の要素と対応するかを話し合え。
- 問 3 また、「ヘッド」の「移動」はどんな機能あるいは行為なのかを話し合え。
- 問 4 言語を抽象化するとはどういうことかを考えなさい。
- 問 5 チューリングマシンの形式的定義 (数理モデル) を調べて、それがどのように表現されているのかを考えなさい。たとえば、[Wikipedia: チューリングマシン](#)にある数式を考えてみよ。

3 マルコフ連鎖

アンドレイ・マルコフ (1856–1922) はロシアの数学者で、確率論において多くの業績を残す^{*2}。次の文を読んで、マルコフ連鎖とチューリングマシンとの関係について考えよ。



Andrei Markov 1856–1922

マルコフ過程とマルコフモデル

ある文字 (あるいは単語) の出現確率が、直前の m 個の記号によって決定されるような確率過程を、「マルコフ過程 (Markov Process)」という。その中でも $m = 1$ の場合、いわゆる直前の場合を特に、単純マルコフ過程と呼ぶ。このマルコフ過程にしたがい、記号が出現すると仮定する確率モデルのことを、「マルコフモデル (Markov Model)」という。 $m = 1$ の、ある時 t の記号の出現確率は、直前の記号のみに依存する。時刻 t における記号を x_t と書くと、マルコフ過程は、 $p(x_t | x_{t-1}, x_{t-2}, x_{t-3}, \dots, x_{t-m})$ となる。直前の記号のみに依存する単純マルコフ過程は、 $p(x_t | x_{t-1})$ となる。マルコフモデルは、[Manning and Schütze \(1999\)](#) の第9章がすべてそれに当てられている。

問6 マルコフモデルとチューリングマシンがどんな点で似ているか、話し合いなさい。

問7 言語 (単語、文、段落など) をマルコフモデルで計算するためにはどのような情報を用いて計算すれば良いかを話し合いなさい。

問8 n-gram モデル (シャノンの情報理論) も隣接の確率を計算するモデルである。n-gram モデルを調べ、どのようなモデルで、どのようなところで利用されているか、調べてみなさい。

4 Markov Chain による文章の生成

Mark V. Shaney とは次のようなものである。

[Mark V. Shaney](#) is a fake Usenet user whose postings were generated by using Markov chain techniques. The name is a play on the words “Markov chain.” Many readers were fooled into thinking that the quirky, sometimes uncannily topical posts were written by a real person.

上記の仕様を再現してみる。単語分割には [KyTea](#) を用い、それを [shaney](#) に渡す。

リスト 1: 「坊っちゃん」をマルコフチェーンで表現する。

```
% head -2000 bochan.txt | kytea -notags | sed -e '{s/。/。 \n/g; s/ //g}' | shaney | tr
-d '\n' | sed -e 's/。/。 \n/g' | sed -e 's/ //g' | head -10 | nl
1 坊っちゃん何時家を御始めなすっては如何ですよ。
2 二時間目も昼過ぎの一時間目も昼過ぎの一時間も大同小異である日三階から威勢よくない。
3 やがて湯に染ったが、帰っている。
4 おれがあきれ返って、あ痛いがいいと云った。
5 冗談も度を過ごせばすぐ出来る。
6 あきれ返った時、おやじが大きな眼を御覧なさいますものをつらまえて、仰山な音がする。
7 夕べは寐られていた。
8 邪魔になって、あさってから学校へ行く考えも何もない恰好である私立の中学校を卒業した。
9 「君の指を切って、翌日学校へ縛りつけて机と睨めっくらをさした。
10 おれはこの手拭が湯に入れてくれと款待なしている連中よりは感心だ。
```

問9 shaney はどのような出力をするのかを考えなさい。

問10 上の出力を見て、これらの文と普通の日本語との違いについて議論せよ。

^{*2} <http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/Markov.html>

問 11 言語を記述するという事は、どういうことかを議論せよ。

5 もっと知るには

- ロボットとチューリングに関わる話、万能チューリング機械については、有本 (2000: 52-62) に書かれている。
- 万能チューリング機械の作り方は有本 (2000: 58-9) に書かれている。ちなみに「ロボットにかける夢」は東工大図書館にかなりおいてあり、そんなに厚くない本なのですぐに読めるだろう。
- 隣接を計算するモデルとして n-gram モデルを紹介したが、オートマトンも隣接のセルに様々な状態を記述し、条件にしたがって、セルの間を移動し、推移の過程を経て、最終的に受理した状態を計算するモデルである。オートマトンが計算機科学の中で、あるいは社会の中（電器製品でも食品でも社会構造でも人体でも何でもいい）で、どのような役割を担っているのかを調べてみよう。
- マルコフモデルは、Manning and Schütze (1999) の第 9 章がすべてそれに当てられている。
- クロード・シャノンが情報理論を創始した論文 “A mathematical theory of communication” で、マルコフ連鎖を利用して、エントロピー（情報量）の概念を説明している。
- 「はじめての AI プログラミング — C 言語で作る人工知能と人工無能」(小高 2006) には、n-gram、マルコフ連鎖についての解説とプログラミング例がついている。C 言語ソースもインターネットで配布されているので、実際に自分で動かしてみることができる。
- マルコフ連鎖ジェネレータが公開されており、それは JavaScript で書かれているので、どのような方法で実装しているかを見ることができる。
- インターネットには Bruce Ellis による Mark V. Shaney の Unofficial Biography がある。
- 上記の URL は移動している可能性もあるので、注意すること。なお、それぞれを QR コードで脚注に掲載しておく*3。
- アラン・チューリングの伝記や業績については、「甦るチューリング—コンピュータ科学に残された夢」(星野 2002) がある。

参考文献

- 有本卓 (2000) 『ロボットにかける夢: 知と身体の調和をめざして』, 岩波科学ライブラリー, 岩波書店, 東京, Japan . アラン・チューリングに関わるページ p.52-63 .
- 星野力 (2002) 『甦るチューリング—コンピュータ科学に残された夢』, NTT 出版 .
- Manning, Christopher D. and Hinrich Schütze (1999) *Foundation of statistical natural language processing*, Cambridge, Massachusetts: The MIT press.
- 小高知宏 (2006) 『はじめての AI プログラミング - C 言語で作る人工知能と人工無能』, オーム社 .



*3 [Unofficial Biography](#).



[JavaScript マルコフ連鎖](#).