



Title	言語学と確率論〔Ⅰ〕：「マルコフ過程」
Author(s)	森塚, 文雄
Citation	大阪外大英米研究. 1983, 13, p. 21-34
Version Type	VoR
URL	https://hdl.handle.net/11094/99061
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

言語学と確率論〔Ⅰ〕

「マルコフ過程」

森 塚 文 雄

1. 自然言語の数学的研究、つまり近代の数理言語学 (Mathematical linguistics) の魁は、実は古くからいろいろの立場で試みられ、とくに人間の思考作用をすべて記号を用いて表現し、これを機械的に処理しようとする考えは、ライプニッツ (Gottfried Wilhelm Leibniz, 1646-1716) 以来多くの先覚の夢であった。

一方では、最近における^{デジタル}数値型万能計算機の急速な発達普及に伴って、プログラム方式の高度の進歩により、人間と機械との間の複雑な内容の通信や、さらには機械をいわば人間の一部として有機的にとり入れた一種の思考過程というようなものが実現されるようになって、言語という通信形式の範囲は急速に拡大され、言語理論も急速な発展を余儀なくされるに至った。ある程度の曖昧さ・漠然性をもった人間の創造的な知的過程である自然言語という離散記号的な通信の枠組を、われわれ人間は先天的に活用し発展せしめる能力を備えていると言われるが、われわれの一人一人は、この先天的な能力によって、一方では他の成員と緊密に協力して社会を形成し、文化の進展を図るが、他方では自己の内部において情報を処理し、記憶し、蓄積し、無限の新しい総合を創造する可能性を与えられている。この「無限の創造性」¹⁾ という特性は、自然言語のみでなく、電信のような簡単な通信形式から電子計算機のような龐大な記憶容量の機械系にも備わっているべきものであって、自然言語の文法記述における繰返し型 (Recursive) 要素の問題も、機械における有限の種類の方法の無限の組合わせの構造の問題も、共通の解法が得られる。²⁾

さて、自然言語はその最も基本的かつ普遍的な形として、音声現象として捉えられるが、これは言語通信の信号としての音声波を空気の振動 (音波) として伝え

る物理現象であって、現象的に「連続量」を取扱う過程とみることができる。しかし観測される現象そのものの**連続性**と、現象によって表記されるべき記号体系 — たとえばヨーロッパ諸言語に広く行われる音素表記による単語の表記 — の**離散性**との間の関係は、直観的に認識され難い傾向があって、そこに言語学が自然科学の進歩に取り残されてきた宿命がうかがわれる。しかしながら、現代の自然科学における高度の理論体系の基礎となった自然法則の発見には、偉大な科学者の直観に基づく仮説の設定に負うところもあるように、人間の知的過程には常に多少の能動的な論理の飛躍がうかがわれる。ただ厳密な意味での科学としての言語学が進歩するためには、その**検証**(Verification)として追求されるべき論理の筋道は、客観的で明確で数学的に厳密に形式化されていなくてはならない。そのような要求をみたす唯一の言語理論が彼の生成変形文法の理論である、と言うのが Noam Chomsky の主張である。つまり、古くからいろいろの立場から試みられた自然言語の数学的研究は、Chomsky によってはじめて厳密に定式化された数理言語学として20世紀の後半に生まれたものと言えるであろう。より形式化された数学理論では、先ず**公理**の集合が定義され、**定理**は推論規則(Rules of inference)を用いて公理から全く形式的に導き出される。そして数理言語学では、上述の定理が**文章**に対応し、推論規則をもつ公理の集合が**文法**に対応し、そのような言語を記述するための形式言語が高等論理言語(**メタ言語**)である。

1) 藤村靖教授「シンタックスの論理」(岩波講座 哲学)

2) 生成文法の理論で文法が扱う最大の言語形式としての単位である「文」は、文法的に環境に依らず独立にその内部構造を決定できると言える最小の単位である。文は、生成過程で繰返し型の単位として働き、1つの文を他の文の中にはめこん(embed)でいくことができ、このような操作を何度でも繰返すことによって理論的には幾らでも大きな文を作ることができるのは衆知の通りである。

2. 以上見てきたところから、Chomsky の生成文法の理論が、それに先立つ構造主義言語学者らやその他の学者の文法理論に対してもつ際だった特性の一つは、これらの理論がみな均しく「科学性」(Scientificallness)を標榜する中

で、Chomsky のそれのみが著しい「数学性」(“Mathematicalness”)をもっているという点であろうと思われる。事実、彼の著書・論文を目にして気づくことは、“Mapping”(写像)，“Boolean(algebra)”(ブール代数)，“Data-processing”(データ処理)，“Embedding”(埋め込み)，“Matrix”(行列)，“Partially ordered”(半順序の)，“Programming”(プログラミング)，“Set”(集合)，“Sub-set”(部分集合)，“Turing machine”(チューリング機械)，“Variable”(変数)など近代数学の重要な術語及びそれらを通して窺われる Chomsky 自身の数学志向性が随所に現われている点である。とりわけ、近代数学の全分野に亘ってその基礎理論となっている「集合論」(Set theory)の考えが Chomsky 理論の根底にはうかがわれる。彼の“Transformation”(変形、変換)や“generate”(生成する)は、それぞれ、集合論、群論の基本概念であることに気づく。

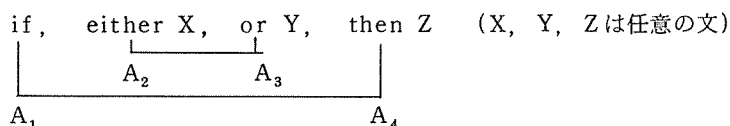
本稿では Chomsky¹⁾ が用いた“(Finite state) Markov process”(有限状態マルコフ過程)を中心にとり上げ、先ず Chomsky 自身がこの術語を彼の文法理論の中でどのように位置づけているかを検討し、次いで情報理論、確率・統計論において如何に扱われているかについて、主に近代情報理論の泰斗 Claude Elwood Shannon(1916—)の、古典とも言われる名著 *The Mathematical Theory of Communication*, '64 に従って比較検討を加えつつ、多少の私見を挿みたい。

Chomsky の“Markov process”, “Markov chain”, “Finite state Markov process”に関する解釈および評価の大略は次のようであろうと思われる。

英語の文法的文を構成する形態素(もしくは語)の連続体のすべて(しかもそれだけ)を生み出す文法に対しては、情報理論による言語のモデルが問題解決の1つの道を示してくれる。ここに1つの機械(体系)を仮定し、それが有限個の“Internal states”(内部状態) S_0, S_1, \dots, S_n の何れか1つをとることができ、ある内部状態 S_i からそれに直接つながる他の状態 S_j に移行する際に、一定の記号(例えば英語の単語)を生み出せるような仕組みになっていて、これを

“Finite state Markov process”(有限状態マルコフ過程)と考える。²⁾ここである内部状態 S_j への“Transition”(遷移)が、その直前の1個の状態 S_i のみに依存していて、それ以前に辿った状態遷移の経路、例えば S_{i-1} , S_{i-2} , S_{i-3} , …… には依存しないところがMarkov processの特徴であって、その意味で“Phrase structure”(句構造)の“Re-writing rules”(書替え規則)はマルコフ過程的であるという。この機械が“Initial state”(最初の状態)から始まって順次一定の記号を生み出しつつ、“Final state”(最後の状態)に達すれば、記号列が得られ、これを文とすれば、その文の集合がその機械によって定義される1つの言語である。このような言語が“Finite state language”であって、その機械(マルコフ過程)は、“Finite state grammar”(有限状態の文法)とすることができる。

而して英語には、例えば



のような“Mirror image”(鏡像型)の依存関係をもつ文があって、このXが更に又鏡像型の依存関係をもつ、というように無限の可能性、すなわち、一般に

$$A_1, A_2, A_3, \dots, A_{n-2}, A_{n-1}, A_n$$

という文の A_1 と A_n , A_2 と A_{n-1} , A_3 と A_{n-2} , …… というような依存関係がある場合には、依存関係とともに状態数も際限がなくなって、「有限状態のマルコフ過程」では、そのような文のみを無限に生み出すことはできない“English is not a finite state language”であってみれば、マルコフ過程的言語観は少なくとも文法には受け入れ難いことを主張している。

1) N. Chomsky : *Syntactic Structures, Aspects*, “Three Models”

2) 数理言語学では、狭義の四則演算にはじまる数多くのタイプの計算の概念の拡張とともに、それを実行する機械的手順である「アルゴリズム」の概念と密接に関係する。アルゴリズム的計算は、離散要素上で行われ、同時にそ

これらの有限個の要素上でのみ実行される。計算のルールは、必ず個数も大きさも有限である。アルゴリズムは、いろいろな論理的目的のために定義されていて、その主なものにチューリング機械(Turing machine)、シューステム(Thue system)、マルコフ・アルゴリズム(Markov algorithm)などがあるが、Chomskyの「有限状態マルコフ過程」による文生成の機械は、その計算原理がチューリング万能機械のそれであろうと考えられる。

3. 以上のように「有限状態のマルコフ過程」はChomskyの生成文法理論では、そのモデルが自然言語を記述するには不適当であるという烙印をおされはしたものの、この「マルコフ過程」は情報理論(Information theory)においてはマルコフ情報源として、最も基本的な重要な概念として取扱われている点に注目したい。Chomskyが念頭に描いていたマルコフ過程は後述する**単純マルコフ過程**もしくは**1重マルコフ過程**のみであったように思われるが、一般に、日本語にせよ英語にせよ、その文章が通信手段として用いられる場合には、「時間とともに変動する偶発現象を数学的に抽象化し定式化された」**確率過程**(Random process)または**ストカスティック過程**(Stochastic process)の1種に過ぎないものであって、自然語にあっては相当多重の確率過程を考える必要があるように思われる。

さてここで前述したShannonがその*The Mathematical Theory of Communication*において**離散的情報源**(Discrete source of information)としてのマルコフ情報源をどのように扱っているかについて、多少の補足的説明や私見を混えつつ、概略の説明をしたいと思う。

5. マルコフ過程とエルゴート過程

よく話題に上る事例で、1枚の貨幣を多数回空中に投げて表または裏が出る確率が各々 $1/2$ であるとか、雙六の骰子の1, …… , 6 の目が出る確率が何れも $1/6$ であるとか、よく混ぜたトランプの1組の札の中から例えばハートのエースを抜き出す確率が $1/52$ であるとか、と言うとき、相互に排他的、独立的である**事象(Event)**が**大数の法則(Law of large number)**のもとで生起するこ

とを意味する。一般にいくつかの可能性のある事象があって、そのどれが実際に生起するかを予め知ることが不可能で、ただ各事象のおきる確率のみがわかっているとき、これを**確率事象**(Random event)といい、この一連の確率事象が時間をパラメータとするとき**確率過程**または**ストカスティック過程**(Stochastic process)と呼ぶ。前述の52枚のトランプの組から1枚のカードを何度も抜き出す場合が確率過程の1例である。この時、同じ試行を何回もくり返す確率過程にあっては、(毎回トランプ全体を余ほどよくかき混ぜない限り)ふつう直前回の試行の影響が最も大きく、前々回の影響も直前回のそれほどではないが、矢張り残っているであろう。しかしこの影響は試行の回数を遡るに従って微弱となり遂に100回前とか1000回前ともなると完全に消滅するであろう。このように、現在ある事象が生起する確率が、以前の試行の結果に影響されるが、ある有限な数まで遡ると、それよりさらに以前の歴史には最早や影響されない。そのような確率過程が**マルコフ過程**である。

今、情報源から発生する事象が

$$\alpha_{i-n-m}, \dots, \alpha_{i-n}, \alpha_{i-n+1}, \dots, \alpha_{i-2}, \alpha_{i-1}$$

の順序になっているとし、条件付き確率分布の間に次の(1)式のような関係があるならば、この情報源は**n重マルコフ情報源**であるという。

$$\begin{aligned} P(\alpha_i | \alpha_{i-1}, \alpha_{i-2}, \dots, \alpha_{i-n}, \dots, \alpha_{i-n-m}) \\ = P(\alpha_i | \alpha_{i-1}, \alpha_{i-2}, \dots, \alpha_{i-n}) \end{aligned} \quad (1)$$

式(1)は、 α_i はさかのぼって α_{i-n} までの事象に依存するけれども、 α_{i-n-1} 以前の事象には無関係であるということを意味する。Chomsky がマルコフ過程として挙げているものは、実はこのマルコフ情報源のうち $n=1$ という特別の場合であって、上の(1)式は

$$P(\alpha_i | \alpha_{i-1}, \alpha_{i-2}, \dots, \alpha_{i-1-m}) = P(\alpha_i | \alpha_{i-1})$$

となり、前回の試行だけに影響される**単純マルコフ過程**である。細かい説明は抜きにして、単純マルコフ過程の確率的な構造は一般に**遷移確率**(**推移確率**)(Transition probability) $P_i(j)$ か、又は2文字の確率 $P(i, j)$ で記述される。

もう一段複雑なマルコフ過程として考えられるものは、遷移確率が先行する2

つの事象に関係し、それ以前の事象には関係しない確率過程であって、これは2重マルコフ過程といい、一般に遷移確率が先行する n 事象までに関係し、それ以前の事象には関係しない確率を n 重マルコフ過程という。1 重マルコフ過程の場合と同様に、2 重マルコフ過程は遷移確率 $P_{ij(k)}$ 、又は3文字の確率 $P(i, j, k)$ で記述され、両者の間には

$$P(i, j, k) = P(i, j)P_{i,j(k)} \quad (2)$$

という関係がある。一般に n 重マルコフ過程は遷移確率 $\underbrace{P_{i,j, \dots, k(l)}}_n$ 、又は $n+1$ 文字の確率 $\underbrace{P(i, j, \dots, k, l)}_{n+1}$ で記述され、両者の間に

$$P(i, j, \dots, k, l) = P(i, j, \dots, k)P_{i,j, \dots, k(l)} \quad (3)$$

という関係がある。更にマルコフ過程の特別な例とみなされるものに、情報源から生起する各事象が1つ1つ独立であって互いに影響を与えない0重マルコフ過程が考えられる。

英語や日本語のような自然言語も数学的に確率過程として取扱うことが勿論可能であって、普通の英文はおよそ8重マルコフ過程の情報源であると言われているが、これは英語の単語を構成する文字数の平均が約4・5であるとされるところから、1語ないし2語前の単語と相関があることになり、情報源としての英文は語という今1つ上位の事象から見て単純もしくは2重マルコフ情報源とみなされるであろう。

以上のマルコフ過程のうちで、とくに情報理論にとって重要なものは、エルゴード過程(Ergodic process)と呼ばれる確率過程であるが、簡単に言うところは一連の試行の結果として得られる文字系列において、大数の法則が成立するようなマルコフ過程である。すなわち、エルゴード的なマルコフ過程では、一連の十分に長い文字系列が得られたとき、大部分の系列では、その中に含まれる1文字の頻度とか、2文字の頻度とか、状態から状態への遷移の頻度を調べると、文字の系列が長くなるほど、それらが1文字の確率とか、2文字の確率とか、遷移確率にほとんど正確に近づき、反対に、それから外れるような系列が得られる

確率は零に限りなく近づく。

6. 離散的情報源

ここで離散的情報源というのは、有限集合から選び出された離散的記号系列を生む確率過程であって、日本語や英語などの自然語の書きことば、PCM送信機の量化音声やテレビの量化信号のような量化過程で離散的となった連続的情報源の外、次のような符号連続を生成する確率過程を抽象的に規定する数学的実例をいう。

- (A) すべて 0.2 の等確率をもち連続した 2 字が独立である A, B, C, D, E の 5 字の乱数表による系列、例えば、

BDCBCECCCADCBDDAAECEEAABBDAAEECACEEBAEE
CBCEAD.

- (B) 各々 0.4, 0.1, 0.2, 0.2, 0.1 の確率をもち連続 2 字が独立である A, B, C, D, E 5 字の系列、¹⁾ 例えば、

AAACDCBDCEAADADACEDAEADCABEDADDCECAAAA
AD.

- (C) 連続記号が相互に独立でなく、確率は先行記号(1語のみのときは“digram”(2字連接), 2語のときは“trigram”(3字連接), 一般に $n-1$ のときは“ n -gram”(n 字連接)) に依存する複雑な仕組 A, B, C 3字のみの 2字連接の例、

ABBABABABABABABBBABBBBBBABABABABABBBACA
CABBABBBBABABACBBBABA.

- (D) 「語」の系列からなるテキスト生成の確率過程、今 A, B, C, D, E の 5 字と次の確率をもつ 16 「語」を有する言語¹⁾を仮定すると、

0.10 A	0.16 BEBE	0.11 CABED
0.04 DEB	0.04 BED	0.05 CEED
0.15 DEED	0.05 ADEE	0.02 BEED
0.08 DAB	0.01 EAB	0.01 BADD
0.05 CA	0.04 DAD	0.05 EE

これから作られる典型的な通報は

DAB EE A BEBE DEED DEB ADEE ADEE EE
DEB BEBE BEBE BEBE ADEE BED DEED DEED
CEED ADEE A DEED DEED BEBE CABED BEBE
BED DAB DEED ADEB. となる。

上述の人工言語は、いろいろの可能性を例示する簡単な問題や実例を構成するのに重宝である。又、一連の単純な人工言語を用いて自然言語への近似を行なうこともできる。すなわち、

- (1) “Zero-order approximation” (ゼロ次近似)では、すべての文字が等確率にかつ独立に選ばれ、
- (2) “First-order approximation” (第1次近似)では、文字連続が独立だが各文字が自然言語と同一の確率をもつように選ばれる。(例えば、英語への第1次近似であれば、Eは通常の英語のEの確率と同じく 0.12の確率を、Wは 0.02の確率をもつが、TH, ED などのように頻度の高い2字接続はなく、相接する2字は相互に独立である。)
- (3) “Second-order approximation” (第2次近似)では、2文字接続が導入され、1文字が選ばれると、後続するいろいろの文字の確率に応じて次の文字が選ばれる。
- (4) “Third-order approximation” (第3次近似)では、3文字接続が導入され、各文字は先行する2文字に依存する確率で選ばれる。

以上の一連の過程が如何ように言語に接近するかを示すために、英語への近似の典型的な諸系列を構成し披露する。すべての場合、英語の26文字と単語スペースとを加えた「27文字アルファベット」を考える。

- (1) “Zero-order approximation” (各記号は独立でかつ等確率)

XFOML RXKHRJFFJUJ ZLPWCFWKCYJ
FFJEYVKCQSGHYD QPAAMKBZAACIBZLHJQD.

これは英語のアルファベットが使っているというだけで、各文字間に相互関係が全くなく、しかも等確率であれば、現実の英語とは全く似ても似つかず、

意味をなさないものである。

- (2) “First-order approximation” (各記号は独立だが、確率は英語のテキストと同じ)

OCRO HLI RGWR NMIELWIS EU LL NBNESEBYA
TH EEI ALHENHTTPA OOBTTVA NAH BRL.

英語のアルファベットの中、もっとも頻度の高い単語スペースやE, T, A, O, Nなどが同じ高頻度で現われはしているが、各文字間の相関々係がないため、矢張り無意味である。只、単語の長さが英語らしくなる。

- (3) “Second-order approximation” (英語と同じ2文字連接)

ON IE ANTSOUTINYS ARE T INCTORE ST BE
S DEAMY ACHIN D ILONASIVE TUCOOWE AT
TEASONARE FUSO TIZIN ANDY TOBE SEACE
CTISBE.

英語に関しては2文字連接の頻度も調べられているが、それに依ると、例えば文字Qの次にUの来る確率は1で、他の文字の来る確率は零であることはもちろんだが、Bの次にEが来てBEとなる頻度は高く、Tが来てBTとなる頻度は低いというようなことが分かっているがこれはわれわれの直観とも一致する。

つまり言語では意味のある文字の配列は、過去にどのような文字が並んでいたかに重要な意味があることがわかる。そこで、言語をマルコフ過程の如くに考えれば、その言語の確率的性格をとらえうることが予想される。この「第2次近似」は単純マルコフ過程と考えるのが妥当であろう。

- (4) “Third-order approximation” (英語と同じ3文字連接)

IN NO IST LAT WHEY CRATICT FROURE BIRS
GROCID PONDENOME OF DEMONSTURES OF THE
REPTAGIN IS REGOACTIONA OF CRE.

ここに来ると、少し英語らしくなっては来たが、まだ意味のある文章にはなっていない。しかし更に多重のマルコフ過程と考えていけば、次第に英語に対するよい近似になることは予想される。

- (5) “First-order word approximation” (4文字連接…… n 文字連接を考えるよりも、もっと近似を高める手とり早いやり方は文字の代りに単語を使うことである。各単語を独立に、確率分布に従って調べる。)

REPRESENTING AND SPEEDILY IS AN GOOD
APT OR COME CAN DIFFERENT NATURAL HERE
HE THE A IN CAME THE TO OF TO EXPERT
GRAY COME TO FURNISHES THE LINE MESSAGE
HAD BE THESE.

- (6) “Second-order word approximation” (単語の遷移確率は正しいが、それ以上の構造は含まれていない。)

THE HEAD AND IN FRONTAL ATTACK ON AN
ENGLISH WRITER THAT THE CHARACTER OF
THIS POINT IS THEREFORE ANOTHER METHOD
FOR THE LETTERS THAT THE TIME OF WHO
EVER TOLD THE PROBLEM FOR AN UNEXPECTED.

以上の各段階を上ぼるに従って通常の英語のテキストへの近似が目覚しく増大することが分かる。(6)の単語による第2近似ではよほど英語らしくなってきた。“attack on an English writer that the character of this”という単語の連続は全然おかしくはない。このようにして、十分長い過去を考慮したマルコフ過程を考えれば、英語をいくらでも正確にあらわしうることが理解される。

以上は英語を例にとって、言語がエルゴートのマルコフ過程として記述できることに言及したが、このことは英語に限らずいづれの言語についても成立するであろうと考えられる。すべての情報源からの情報の発生が、エルゴートのマルコフ過程として記述できるものと仮定して、その仮定によって近代情報理論はその数学的な基礎が与えられ、その基礎の上に理論を発展させていくことができるようになったのである。実にこのマルコフ過程の情報理論への導入は、1文字当たりの情報量としてのエントロピー(Entropy)や、所与の情報源がどれだけ無

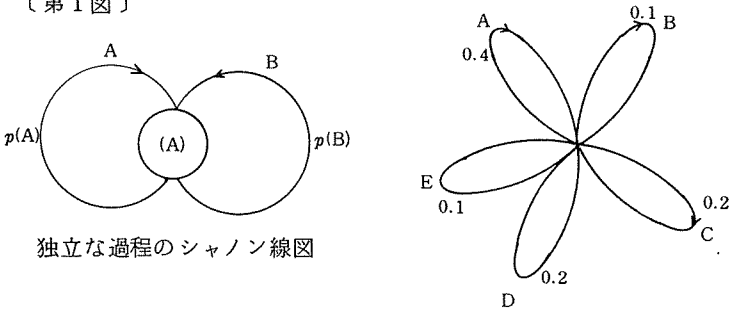
駄なものを含んでいるかの度合を示す冗長度 (Redundancy) などとともに Shannon 情報理論の根幹をなす重要な概念である。これらについては別の機会を待つこととしたい。

われわれは西洋書の一頁を漫然と開いて、しばしば目に入る文字の結合に注目すれば、たとえその単語の意味や文全体の意味が分からなくても、既知の言語であるか否か、既知の言語ならば、どの言語なのかが、およそ見当がつくことが多い。例えば、sch, tz, dt などの2文字〔3文字〕接続からはドイツ語が、oi, eau, eur などの2文字〔3文字〕接続からはフランス語が、er, ing, str, spr などの2文字〔3文字〕接続からは英語が、ほぼ推察されるというように、特定の文字接続の異常に高い頻度や接続する文字間の高い依存度 (Dependency) などが各自然語の書記構造を特色づけるものであってみれば、これらの問題に深くかかわるマルコフ過程が情報理論からの新しい言語研究に如何ほど重要であるかが十分察せられる。

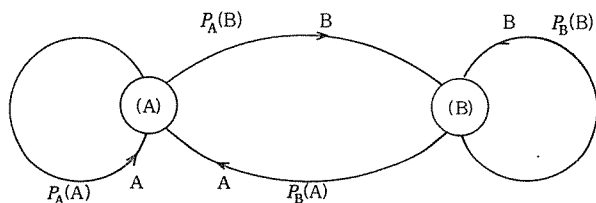
— 未完 —

1) シャノン線図 (Shannon diagram) によれば、マルコフ過程は現在の状態を表わす丸印と、その状態で確率事象がおきることを表わす矢印とで表現できる。そこで、ある事象が起きるということは状態から状態への遷移 (Transition) と考えられ、マルコフ過程の確率的な構造は、状態から状態への遷移確率によって記述される。次に独立な過程のシャノン線図〔第1図〕、単純マルコフ過程のシャノン線図〔第2図〕、および、ここであげた人工言語を構成する単語系を文字の遷移確率で表わしたシャノン線図を示そう。

〔第1図〕

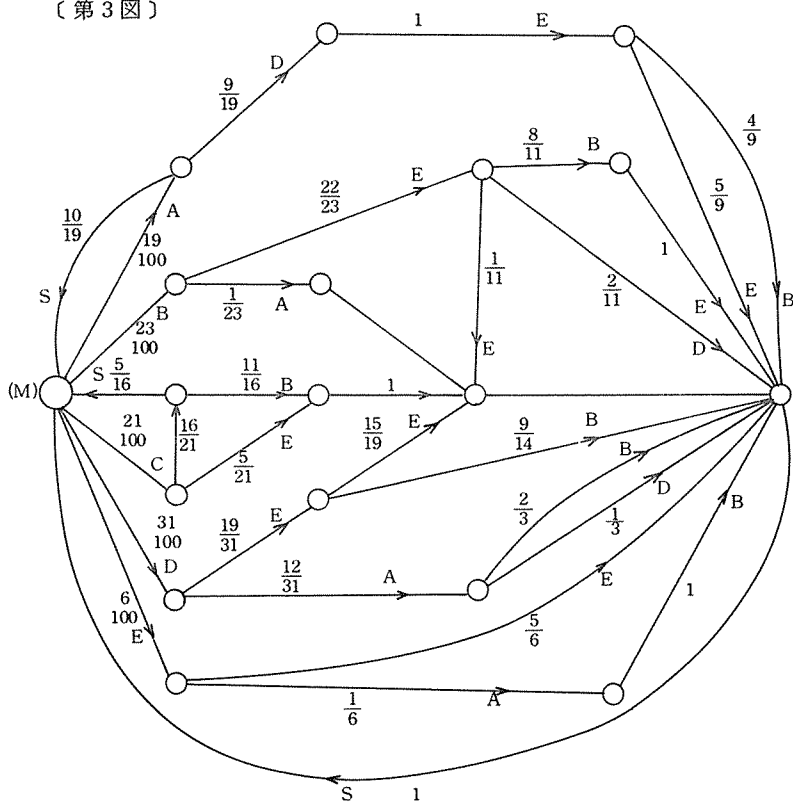


〔第2図〕



単純マルコフ過程のシャノン線図

〔第3図〕



単語系を文字の遷移確率で表わしたシャノン線図

主な参考書目

- Claude E. Shannon and Warren Weaver: *The Mathematical Theory of Communication*, The University of Illinois Press, Urbana, 1964.
- Colin Cherry: *On Human Communication, A Review, A Survey, and A Criticism*, 3rd Edition, The M.I.T. Press, 1977.
- Noam Chomsky: *Syntactic Structures*, Mouton, 1st Printing 1957.
- Noam Chomsky: *Aspects of the Theory of Syntax*, The M.I.T. Press, 1965.
- Noam Chomsky: "Three Models for the Description of Language" (I.R.E. Transactions on Information Theory, IT-2)
- Noam Chomsky: 'On the Notion "Rule of Grammar",' (Proceedings of Symposia in Applied Mathematics, Volume XII), *Structure of Language and Its Mathematical Aspects*, American Mathematical Society, 1961.
- S.K. Shaumyan: *Applied Grammar as a Semantic Theory of Natural Language*, Edinburgh University Press, 1977.
- Robert de Beaugrande and Wolfgang Dresser: *Introduction to Text Linguistics*, Longman, 1972.

- 近藤・好並, 「論理学概論」(岩波新書), 1964.
- 栗原・吉田, 「言語理論」(情報科学講座, C. 10・1)(共立出版)(1966).
- 前原昭二, 「記号論理読本」(日本評論社), 1976.
- 関英男, 「情報理論」(オーム社), 1969.
- 本多波雄, 「情報理論入門」(日刊工業新聞社), 1960.
- 岩波講座「哲学 — 言語」, 1971.
- 野崎昭弘, 「電子計算機と数学」(ダイヤモンド社), 1976.
- 岩波「数学辞典」第2版(岩波書店), 1968.