# 言語の認知科学配布資料
# 形式言語

浅川伸一

2009 年 11 月 24 日

## 1　言語の統計

言語の中でもっとも簡単に調べられるのは頻度（何回出現したか）データである。図 1 に「不思議の国のアリス」に出現する文字の頻度を示した。表

表 1: 「不思議の国のアリス」に出現する各文字の頻度

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| SP | 24639 | d | 4931 | f | 2001 | ; | 194 |
| e | 13575 | l | 4716 | p | 1524 | x | 148 |
| t | 10688 | u | 3468 | b | 1475 | j | 146 |
| a | 8791 | ’ | 2871 | k | 1158 | ” | 113 |
| o | 8146 | w | 2675 | . | 989 | z | 78 |
| i | 7514 | g | 2531 | v | 846 | * | 60 |
| h | 7374 | , | 2418 | ! | 450 | ) | 56 |
| n | 7016 | c | 2399 | & | 233 | ( | 56 |
| s | 6500 | y | 2262 | q | 209 | | |
| r | 5438 | m | 2107 | ? | 202 | | |

中の SP は空白を表している。図から明らかなとおりもっとも出現頻度の高い文字は空白であり，その次が e である。

英語の文字列の特徴を表すものとして，単独文字の出現頻度だけではなく，2 文字，3 文字が隣接して生じる頻度を問題にする。これは共起関係と呼ばれ，2-gram, 3-gram などと表す。同じく「不思議の国のアリス」から 2-gram, 3-gram を調べたものを以下に記す。この表は，t の後に e が来た回数が 755 回であったことを表している。おのおのの n に対して n-gram を作成すれば，英単語の綴りの特徴が見えてくると考えて良いであろう。

一方，日本語の場合は文字数が 8000 文字を越えることから 2 次の相関表を作るだけでも大変である。宮沢賢治の「銀河鉄道の夜」から文節区切りソ

表 2: 「不思議の国のアリス」にみられる 2 次の共起関係

| | SP | e | t | a | o |
|---|---|---|---|---|---|
| SP | | 337 | 3957 | 3005 | 1279 |
| e | 4487 | 484 | 348 | 775 | 44 |
| t | 2578 | 755 | 364 | 259 | 1008 |
| a | 616 | | 1167 | | 3 |
| o | 1152 | 48 | 425 | 25 | 445 |
| i | 383 | 191 | 1325 | 32 | 174 |

フト mecab を使って一単語を切り出し，そのの頻度をとってきたのが，表 3 である。

表 3: 「銀河鉄道の夜」にみられる単語の頻度情報

| | | | | | |
|---|---|---|---|---|---|
| の | 1276 | が | 533 | ジョバンニ | 189 |
| 。 | 1120 | まし | 451 | い | 185 |
| 、 | 987 | と | 309 | な | 181 |
| た | 948 | 」 | 293 | | 181 |
| て | 866 | 「 | 293 | から | 166 |
| に | 780 | も | 237 | です | 161 |
| 》 | 693 | で | 215 | だ | 157 |
| 《 | 693 | よう | 209 | その | 156 |
| は | 625 | し | 209 | ん | 144 |
| を | 565 | か | 205 | へ | 126 |

表 4 では，太宰治「人間失格」の中から，2 次の相関の高い単語の組み合わせを表示したものである。この表をみると，(は,、) の組み合わせが一番多く 1.51% を占めていることが分かる。日本語の格助詞で主格を表す格助詞の後に読点（、）が来ることが多いことが分かる。

## 1.1 マルコフモデル

一番簡単な言語モデルとしてマルコフモデル Markov model を紹介する。単語の列 $W = w_1 w_2 \cdots wn$ の同時確率 $P(W)$ は次の条件付き確率として書くことができる。

$$P(W) = \prod_{i=1}^{n} P(w_i | w_1 w_2 \cdots w_{i-1}) \tag{1}$$

表 4: 「人間失格」にみられる単語の共起頻度情報 (%)

| | | | | | |
|---|---|---|---|---|---|
| は,、 | 725 (1.51) | て, い | 226 (0.47) | で,、 | 153 (0.31) |
| た,。 | 591 (1.23) | も,、 | 219 (0.45) | です,。 | 148 (0.30) |
| て,、 | 406 (0.84) | て, いる | 214 (0.44) | 」,「 | 131 (0.27) |
| まし, た | 365 (0.76) | に,、 | 191 (0.39) | 、, それ | 123 (0.25) |
| 、, 自分 | 363 (0.75) | 、, その | 190 (0.39) | と,、 | 120 (0.25) |
| でし, た | 313 (0.65) | 自分, の | 178 (0.37) | の, でし | 116 (0.24) |
| 自分, は | 290 (0.60) | た, の | 176 (0.36) | 。,「 | 111 (0.23) |
| が,、 | 279 (0.58 | に, は | 169 (0.35) | 自分, に | 107 (0.22) |
| し, て | 278 (0.57 | 」, | 166 (0.34) | ませ, ん | 104 (0.21) |
| 。, | 268 (0.55 | の, です | 157 (0.32) | よう, な | 98 (0.20) |

さまざまな単語の組み合わせに対して条件付き確率 $P(w_i|w_1w_2\cdots w_{i-1})$ を推定することは現実的には不可能なので，これを $N-1$ 重マルコフ過程で近似したモデルを n—gram モデルという。$N=1$ の場合を unigram, $N=2$ を bigram, $N=3$ を trigram と呼ぶ。単語 bigram の場合は以下の式で表せる。

$$P(W) = \prod_{i=1}^{n} P(w_i|w_{i-1}) \tag{2}$$

右辺の確率は,

$$P(w_i|w_{i-1}) = \frac{c(w_{i-1}, w_i)}{c(w_{i-1})} \tag{3}$$

ここで $C(\cdot)$ は単語列の出現頻度を表す。

## 2 マルコフ過程 Markov process

時刻 $t$ にとる値の分布が過去の ($t$ 以前の) 以前任意の 1 時刻 $t_0$ にとった値だけに関係し，$t_0$ 以前の履歴には影響されない確率過程をいう。正確には，確率過程 $X_{t(t\in T)}$ において，$T$ 内の時刻 $s_1 < s_2 < s_3 < \cdots < s_n < t$ を任意に選んだとき，$\boldsymbol{X}_{s1}, \boldsymbol{X}_{s2}, \cdots, \boldsymbol{X}_{sn}$ を与えたときの $X_t$ の条件付き確率が，$\boldsymbol{X}_{sn}$ だけをあたえたときの $X_t$ の条件付き確率に等しいとき，$\{X_t\}(t \in T)$ をマルコフ過程という。マルコフ過程が与えられたとき，時刻 $s$ で $X_s$ の値をあたえたきに，時刻 $t(\geq s)$ における $X_t$ が領域 $A$ に落ちる条件付き確率を推移確率 transition probability $P(s, x; t, A)(x \in S, A \in B(S), S$ は $X_t$ の値域といい，正確には次のように定義する。

1. $P(s, x; t, A)$ は $x$ について $B(S)$ 可測， A については確率分布である。
2. $P(s, x; s, A) = 1(x \in A), 0(x \notin A)$

3. 確率 1 で $P(X_t \in A|X_s = x_s) = P(s, x_s; t, A)$.

推移確率は確率保存の関係式であるチャップマン･コルモゴロフ方程式 Chapman–Kolmogorov equation

$$P(s, x, \mu, A) = \int_s P(s, x; t, dy)P(t, y, \mu, A) \tag{4}$$

を，$x$ について $X_\gamma$ の分布に関し測度 0 の点を除いて満足する。逆にこの方程式と上の 3 条件を満たす関数 $P(s, x; t, A)$ があたえられ，初期時刻 $t = 0$ における $X_0$ の分布 $F$ が任意に指定されたとき，$P(s, x; t, A)$ を推移確率とし，$F$ を初期分布 $(\Phi_{X0})$ とするマルコフ過程 $\{X_t\}_{(t\in T)}$ が存在する。推移確率が $t$ と $s$ との差だけに関係して，$P(s, x; t, A) = P(t - s, x; A)$ と表せるとき，このマルコフ過程は時間的に一様であるという。また実数値（または $R^n$ 値）マルコフ過程において，推移確率が位置のずれによって変わらないとき，すなわち集合 $A$ を $x$ だけずらせたもの $\{y - x|y \in A\}$ を $A - x$ と書いて，$P(s, x; t, A) = P(s, t; A - x)$ で表せるとき，このマルコフ過程は空間的に一様であるという。空間的に一様なマルコフ過程は加法過程であり，加法過程 $\{X_t - X_0\}_{(t\in T)}$ を初期値 $X_0$ だけずらせた $\{X_t - X_0\}_{(t\in T)}$ は空間的に一様なマルコフ過程である。$X_t$ の値域 $S$ が加算集合であるマルコフ過程をマルコフ連鎖という。また確率 1 で見本過程が連続なマルコフ過程を拡散過程という。

## 2.1 マルコフ連鎖

マルコフ過程 $\{X_t\}_{t\in T}$ のとる値が有限個または加算個のときマルコフ連鎖という。ときに離散的な時間助変数のマルコフ過程をマルコフ連鎖ということもある。以下，時間的に一様な場合を考える。$X_t$ の値域を $S$ とする。$X_s = X$ のとき $t > 0$ に対し $X_{s+t} = y$ となる条件付き確率

$$p_t(x, y) = P(X_{s+t} = y|X_s = x) \tag{5}$$

をマルコフ連鎖の推移確率 transition probability という。これは次の 2 条件を満たす。

$$P_t(x, y) \geq 0, \sum_{y\in S} p_t(x, y) = 1, \tag{6}$$

$$p_{t+s}(x, y) = \sum_{z\in S} p_t(x, z)p_s(z, y). \tag{7}$$

式 (7) をチャップマン・コルモゴロフ方程式という。逆にこの 2 条件を満たす $\{p_t(x, y)\}$ があれば，これを推移確率とするマルコフ連鎖が存在し，初期分布を与えればただ 1 つに定まる。とくに確率行列 $\boldsymbol{A} = (a(x, y))$, すなわち $a(x, y) \geq 0$, $\sum_{y\in S} a(x, y) = 1$ となる行列があるとき，$p_n(x, y) = \boldsymbol{A}^n(n =$

$1, 2, \cdots$) とおけば，$p_n(x, y)$ は上の 2 条件をみたし，これを推移確率とする離散時間助変数のマルコフ連鎖が存在する。$X_t$ が $N$ 次元格子点上を動き，$a(x, y)$ が $y - x$ だけの関数 $a(x, y) = b(y - x)$ となるとき，これから構成されるマルコフ連鎖をランダムウォークという。

岩波，理化学事典より

# 3　言語の有限オートマトン理論

言語の形式的モデルにはいくつかある。そのうち有限オートマトン理論は最も単純なものであり，マルコフモデルの延長上にあると考えて良い [?]。

## 3.1　有限オートマトン理論の基礎概念

言語の書記素 orthography の基本単位は文字である。そこである言語に現れるすべての文字集合を，一般性を失うことなく，**アルファベット** と名付け $S$ で表す。英語の場合には，$S_E = \{a, b, \cdots, z\}$ であり，日本語の場合には，常用漢字 1945 文字だとすれば，$S_J = \{$ 亜, 哀, $\cdots$, 湾, 腕 $\}$ である。

アルファベット $S$ 上で文字列を定義する。これは，$S$ の文字を任意の回数だけ繰り返して並べたものである。アルファベット 1 文字の集合を $S$ とし，2 文字からなる集合を $S^2$, $n$ 個ならべたものを $S^n$ とする。特別なものとして，文字を 0 個並べたものを

$$S^0 = \{\lambda\} \tag{8}$$

と表現する。$\lambda$ を**空語（空列）null string** という。すべての集合の和集合

$$S^* = S^0 + S^1 + S^2 + \cdots + S^n + \cdots = \sum_{i=0}^{\infty} S^i \tag{9}$$

$$S^+ = S^1 + S^2 + \cdots + S^n + \cdots = \sum_{i=1}^{\infty} S^i \tag{10}$$

と表す。$S^*$ をアルファベットの**閉包 closure** と呼ぶ。すなわちアルファベットの文字を任意の個数（0 も含んで）だけ並べた**文字列 character string** 全体の集合である。たとえば $S = \{a, b, c\}$ とすると，$S^0 = \{\lambda\}$, $S^1 = \{a, b, c\}$, $S^2 = \{aa, ab, ac, ba, bb, bc, ca, cb, cc\}$, $S^3 = \{aaa, aab, aac, aba, \cdots\}$, などとなる。アルファベット $S$ を $A \sim Z$, $a \sim z$, $0 \sim 9$, ピリオドやスラッシュなどの記号などにとれば，$S^*$ には英語のすべての単語，すべての文，すべての文章，になる。$S^*$ の中にはシェイクスピアの作品から，これまでに書かれた，そしてこれから書かれるであろうすべての文書を含む集合である。

$S^*$ はアルファベット $S$ のすべての文字列を含んでいるため，特別な興味はない。$S^*$ の中のある文字列 $\alpha$ を取り出してみると，英語の文になっていることもあれば，でたらめな文字列であることもある。ある文字列 $\alpha$ が言語 $L$ の文となっている場合，これを

$$\alpha \in L \tag{11}$$

と書き，そのようなすべての $\alpha$ の集合を**言語** $L$ と定義する。当然ながら，

$$L \subset S^* \tag{12}$$

$L$ は $S^*$ のサブセットであるという。

## 3.2 有限オートマトン

マルコフモデルは，文の生じやすさを確率によって表現したものということができる。ある言語の文として許されるもの，発生しうるものはすべて同等に取り扱うという立場に立てば，マルコフモデルでの確率という概念のかわりに，可能かどうかという 2 値の確率をもったモデルを作ることになる。このようなモデルを**有限オートマトンモデル finite automaton model:fa** と呼ぶ。有限オートマトンは文法によって定義されるすべての言語を定義することはできない。有限オートマトンによって定義しうる言語は正規言語 regular language と呼ばれる。

アルファベット $\Sigma$ 上の有限オートマトン $M$ とは次のような体系 $(K, \Sigma, \delta, q_0, F)$ のことである。ここに $K$ は状態 state の空でない有限集合，$\Sigma$ は有限の入力アルファベット input alphabet, $\delta$ は $K \times \Sigma$[1] から $K$ の中への関数。$q_0$ は初期状態 initial state で $K$ に属し，$F$ は最終状態 final state の集合で $F \subseteq K$ である。

有限オートマトンモデルで，ある与えられた文字列がある言語の文となっているかどうかを検定したいという場合を考える。その文字列をモデルの**初期状態 initial state** に対して与えると，入ってくる文字列の各文字と内部状態の組み合わせで，次に移るべき状態が決定され，有限オートマトンの**状態推移 state trasition** が行われる。文字列を最後まで入れ終わった状態でオートマトンがあらかじめ定められている**最終状態 final state** となっている場合，この文字列は**受理 accept** されたという。このようにして文字列がその言語に属するかどうかが判別できる。このようなオートマトンは次のように定式化でき，

1. 状態の集合を $K = \{q_0, q_1, \cdots, q_n\}$ とする。$q_i$ は状態で，$n$ は有限である。

[1] $K \times \Sigma = \{k, \sigma | k \in K$ かつ $\sigma \in \Sigma\}$

2. 入力文字列の集合を $\Sigma$ とする。$\Sigma$ は有限である。

3. 状態推移関数の集合を $\delta$ とする。$\delta$ は有限である。

4. 初期状態を $q_0$ とする。$q_0 \in K$ である。

5. 最終状態の集合を $F$ とする。$F \subset K$ である。

このような有限オートマトンを

$$M = (K, \Sigma, \delta, q_0, F) \tag{13}$$

と表記する。

文 $x$ は $\delta(q_0, x) = p$ なる $p$ が最終状態の集合 $F$ に属するとき，オートマトン $M$ によって受理されたという。オートマトン $M$ によって受理されるすべての文 $x$ の集合を $T(M)$ であらわす。すなわち，

$$T(M) = \{x | \delta(q_0, x) \text{ は } F \text{ に属する } \} \tag{14}$$

オートマトン $M$ がある状態 $q_i$ にあるとき，入力 $a$ が与えられたら，状態 $q_i$ に移行する場合，

$$\delta(q_i, a) = q_j \tag{15}$$

と書き，**状態推移関数**という。$\delta$ はこのような関数の集合である。この状態推移関数に対応して，次の図 1 のように書ける。

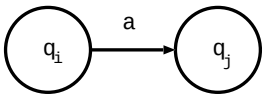


図 1: 状態遷移図

簡単のため，文字 $a$ が任意個並んだ文字を受理する有限オートマトンを作ってみる。これは

$$S^* = \lambda + a + aa + aaa + \cdots + a^n + \cdots \tag{16}$$

という集合を受理するものである。これは図 2 に示すように これは 1 つの状態をもつオートマトンで実現できる。これを詳しく書くと，

$$\begin{array}{lcl} K & = & \{q_0\} \\ \Sigma & = & \{a\} \\ P & = & \{\delta(q_0, a) = q_0\} \\ \text{初期状態} & = & \text{最終状態} = q_0 \end{array} \tag{17}$$

図 3 には，ある有限オートマトンの完全な記述を示した。
$M = (K, \Sigma, \delta, q_0, F)$

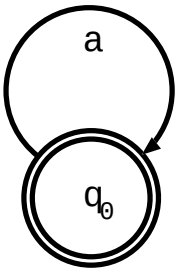


図 2: 状態遷移図その 2

$\Sigma = \{0, 1\}$
$K = \{q_0, q_1, q_2, q_3\}$
$F = \{q_0\}$
$\delta(q_0, 0) = q_2,\ \delta(q_0, 1) = q_1$
$\delta(q_1, 0) = q_3,\ \delta(q_1, 1) = q_0$
$\delta(q_2, 0) = q_0,\ \delta(q_2, 1) = q_3$
$\delta(q_3, 0) = q_1,\ \delta(q_3, 1) = q_2$
状態図は，状態を表すノードと，状態の変化を表す矢印とから成っている。

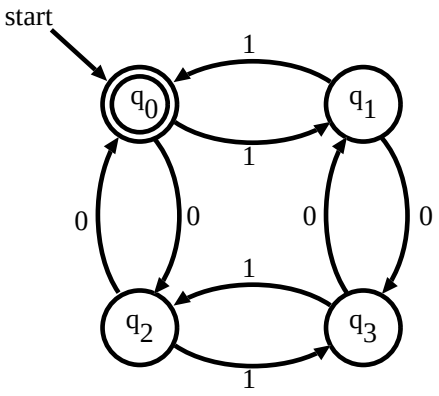


図 3: 偶数個の 0 と偶数個の 1 とをもつ列の集合を受理する有限オートマトンの状態図

すなわち，オートマトンが状態 $q$ で入力記号 $a$ を読んだとき状態 $p$ に移るならば，$q$ から $p$ への矢印を書き，それに標識 $a$ をつける。最終状態，すなわち $F$ に属する状態は 2 重の円で示されている。初期状態は start と書かれた矢印で示されている。

図 3 で $M$ への入力が 110101 であったとする。$\delta(q_0, 1) = q_1$ でしかも $\delta(q_1, 1) = q_0$ であるから，$\delta(q_0, 11) = q_0$. このように $11 \in T(M)$ である。今

問題にしているのは 110101 であるが，$\delta(q_0,0)=q_2$ から $\delta(q_0,110)=q_2$ を得る。次は $\delta(q_2,1)=q_3$ から $\delta(q_0,1101)=q_3$ がわかる。最後に $\delta(q_3,0)=q_1$ および $\delta(q_1,1)=q_0$ から $\delta(q_0,110101)=q_0$　となり，110101 が $T(M)$ に属することがわかる。$T(M)$ が $\{0,1\}^*$ の文で，偶数個の 0 と偶数個の 1 とを含むものの全体の集合であることは容易に確かめられる。

# 4　文法の形式的概念

チョムスキーは句構造文法の考え方を提唱し，文法理論の考えを押し広めた。ここで文法の形式的な概念を整理しておく。

文法には，名詞，動詞句，文などの統語範疇があり，そこから単語の列が導きだされた。これに対応して，統語範疇を非終端記号 nonterminal symbol または変数 variable と呼び，個々の単語を終端記号 terminal symbol と呼ぶことにする。

次の，変数や記号の間の関係を生成規則 production と呼ぶことにする。名詞句 NP → 形容詞 Adj 名詞 N などのように。ある特別な非終端記号があって，その言語に属すると考えられるすべての終端記号列は，その非終端記号から生成される。この特別な非終端記号のことを文，または開始記号と呼ぶ。

形式的には文法 G を $(V_N,V_T,P,S)$ ここで $V_N$ は非終端記号（変数）の集合，$V_T$ は終端記号の集合，$P$ は生成規則の集合，S は開始記号である。$V_N$, $V_T$, $P$ はどれも有限集合である。非終端記号 $V_N$ と終端記号 $V_T$ とをあわせて $V_N \cup V_T = V$ とおく。

空文 empty sentence $\epsilon$ とは 0 個の記号列をさす。アルファベット $V$ 上のすべての文の集合を $V^*$ であらわす。また集合 $V^*-\{\epsilon\}$ を $V^+$ で表す。

生成規則の集合 $P$ は $\alpha \to \beta$ という形の式の集まりである。ここで $\alpha$ は $V^+$ の中の列であり，$\beta$ は $V^*$ の中の列である。

$\alpha \to \beta$ が生成規則 $P$ に属しており，$\gamma$ と $\delta$ とが全ての文 $V^*$ の中の任意の列である場合に $\gamma\alpha\delta \underset{G}{\Longrightarrow} \gamma\beta\delta$ である。このとき $\gamma\alpha\delta$ は文法 $G$ において直接に $\gamma\beta\delta$ を導くという。また生成規則 $\alpha \to \beta$ が列 $\gamma\alpha\delta$ に適用され，列 $\gamma\beta\delta$ が得られたともいう。このように $\underset{G}{\Longrightarrow}$ は 2 つの列を第 2 の列が第 1 の列から一個の生成規則の適用によって得られる場合をさす。$\alpha_1,\alpha_2,\cdots,\alpha_{m-1} \underset{G}{\Longrightarrow} \alpha_m$ であるとする。このとき，$\alpha_1 \overset{*}{\underset{G}{\Longrightarrow}} \alpha_m$ と書く。すなわち 2 つの列 $\alpha$, $\beta$ が $\alpha \overset{*}{\underset{G}{\Longrightarrow}} \beta$ であるとは，$\alpha$ に生成規則 $P$ を何回か適用して $\beta$ が得られることをいう。

文法 G によって生成される言語を

$$\{w | w \text{ は } V_T^* \text{ に属し，かつ } S \overset{*}{\underset{G}{\Longrightarrow}} w\} \tag{18}$$

として定義し，$L(G)$ であらわす。いいかえれば，ある列が $L(G)$ に属するのは次の条件が満たされた場合である。

1. その列は，終端記号のみからなる

2. その列は，$S$ から導き得る

# 5　文法の型 チョムスキー階層

書き換え規則の一般形は，

$$\alpha \to \beta,\ (\alpha \in V^+, \beta \in V^*, V = V_N \cup V_T) \tag{19}$$

と書くことができる。ただし，ここで，$\alpha$ には非終端記号の集合 $V_N$ 中の記号がすくなくとも一つ含まれていなければならない。上式 19 は記号列 $\alpha$ が文字列中に存在すれば，これを記号列 $\beta$ に置き換えることを意味する。そして，$\alpha$ は空列を含まない，その言語に属するすべての文字列の要素であり，$\beta$ は空列を含んだ全ての文字列からなる集合である。

### 5.0.1　0 型文法

初期記号 $\sigma$ から，この形の生成規則を順次適用していって生成される言語 $L(G)$ を **0 型言語** という。この場合の文法を 0 型文法と呼ぶ。

### 5.0.2　1 型文法

式 19 に

$$|\alpha| \leq \|\beta| \tag{20}$$

という制限，すなわち，初期記号の要素の方が書き換え規則を適用した後の要素よりも少ない，という制限をつけた文法を **1 型文法**, あるいは**文脈規定定型句構造文法**, **文脈依存文法**と呼ぶ。

ある文法 $G = (V_N, V_T, P, S)$ で，$P$ のすべての生成規則 $\alpha \to \beta$ が条件 $|\alpha| \leq |\beta|$ をみたすとき，この文法 $G$ を 1 型文法または文脈依存 context sensitive 文法という。文脈依存文法では，ある $\alpha_1$, $\alpha_2$, $\beta \in V^*$, $\beta \neq \epsilon$ および $A \in V_N$ によって $\alpha_1 A \alpha_2 \to \alpha_1 \beta \alpha_2$ のようにあらわされることを要求する。このように制限しても言語の範囲は変わらないことが証明できる。これが文脈依存文法の由来である。

### 5.0.3　2 型文法

さらに生成規則，式 19 が以下のように

$$A \to B,\ (A \in V_N, \beta \in V^+) \tag{21}$$

すなわち $A$ から $\beta$ への書き換え規則のうち $A$ は非終端記号の集合 $V_N$ の要素であり，$\beta$ は，その言語で生成されるすべての文の集合 (より正確にはそこから空列を除いた集合) $V^+$ の要素であるという制約を加えた文法のことを **2 型文法**, あるいは**文脈自由型句構造文法**と呼ぶ。

ある文法 $G = (V_N, V_T, P, S)$ の生成規則 $\alpha \to \beta$ について，

1. $\alpha$ は一個の変数であり，
2. $\beta$ は 空文 $\epsilon$ 以外の任意の列である

が満たされている場合，この文法は 2 型文法または文脈自由 context free 文法であると言われる。$A \to \beta$ という生成規則は，$A$ がどのような文脈に現れるかに無関係に，$A$ を $\beta$ に置き換えることを許すものである。そのため文脈自由という呼び名が生まれた。

#### 5.0.4 3 型文法

文脈自由文法をさらに制約して

$$A \quad \to \quad \alpha B \;\; A, B \in V_N \tag{22}$$

$$A \quad \to \quad \alpha \;\; \alpha \in V_T \tag{23}$$

と制限した場合を **3 型文法**, あるいは**正規文法** と呼ぶ。すなわち $A$, $B$ ともに非終端記号の集合 $V_N$ の要素であり，$\alpha$ は終端記号である場合である。

ある文法 $G = (V_N, V_T, P, S)$ において生成規則 $P$ が $A \to aB$, または $A \to a$ という形であるとする。ここで $A$, $B$ は変数を，$a$ は終端記号をあらわす。このような文法 $G$ は 3 型または正規 regular 文法と呼ばれる。

#### 5.0.5 それぞれの文法の能力

定義から明らかにどの正規文法も文脈自由であり，どの文脈自由文法も文脈依存であり，どの文脈依存文法も 0 型である。文脈依存 context sensitive 文法を csg, 文脈自由 context free 文法を cfg, 正規 regular 文法を rg と略記する。0 型言語は recursively enumerable(再帰的に可算) r.e. 集合と略記される。

0 型文法から 3 型文法までのそれぞれの文法は本質的に異なった能力を持っているとされる。図に示すと図 4 のようになる。ここで，RL は正規言語，CFL は文脈自由形言語，CSL は文脈規定型言語，TM はチューリング機械である。

3 型言語を受理するのは有限オートマトンであり，2 型言語を受理するのは**プッシュダウンオートマトン pushdown automaton** と呼ばれる。1 型言語を受理するモデルは**線形拘束オートマトン linear bounded**

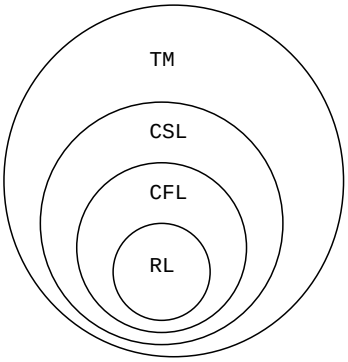


図 4: 形式言語相互間の関係

**automaton** であり，0 型言語を受理するモデルは**チューリング機械 Turing machine** と呼ばれている。