

トランスポゾンから導かれる演算により定義される整合括弧列の DNA モデルとチューリング・マシンの構成

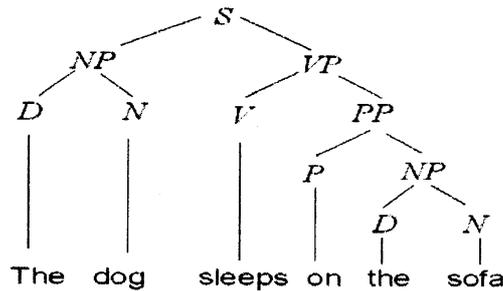
日本大学文理学部情報システム解析学科
鈴木 理

Abstract

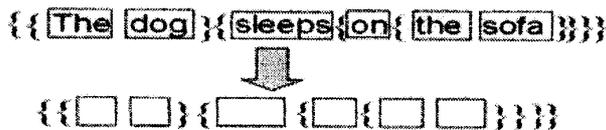
整合括弧列言語に付随した DNA モデルを考える。このときトランスポゾンから定義される演算(T-operation という)、mutation および transition mutation によりこの言語が生成されることが示される。またこの演算を用いて Turing machine が定義されることを述べる。

1. 整合括弧列のつくる言語

チョムスキーは文章の背景には木構造があることを基本としてチョムスキー生成文法理論を構成した([7])。その最も基本的な部分はこの木構造に対して括弧列を考えることにより文脈自由文法(整合括弧列文法)を定義するところにある。英文で書かれた文章の場合の木構造の一例をのべる。

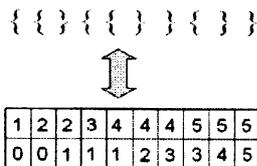


上記の例に対する括弧列の対応はつぎのようになる：

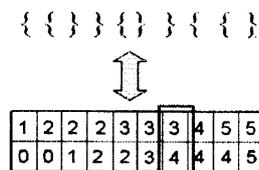


つぎにどのような括弧列が受理言語となるかを判定する。受理言語は括弧列についての TEX-closed 条件により与えられると言える。より正確には次の判定法により与えられる：各段階までの開き括弧の総和を上段に書きその段階まで閉じ括弧列の総和を下段に書く(下図)。上段の総和数が下段の総和数より小さくなく最後にこの両者が等しいとき、またこのときに限り受理される。例をのべる：

Acceptable words



Non-acceptable words



2. 整合括弧列文の分類

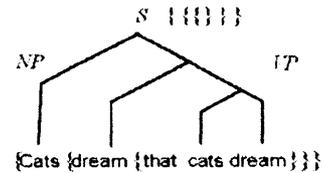
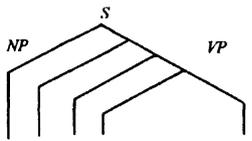
次に整合括弧列文の分類を行う。このためにクラスター数を導入する。括弧列 S をとり開き括弧にひきつづいて閉じ括弧があらわれるときこの間に丸印をかく(下図左)。この丸印の総数を文章のクラスター数といい、 $clus S$ とかく。この数により文章の複雑さが分類される。実際、primary sentence は cluster 1 である(下図上)。cluster 数の高い文章は一般に幾つかの短文が接続詞等で結合されたより高等な文章を記述していると言える(下図下)。

単純文章 (I-型文章)

$clus\{\{\{\}\}\} = 1$ の文章を I-型文章という。

例 $\{\circ\}, \{\{\circ\}\}, \{\{\{\circ\}\}\}, \dots$

Tree structure

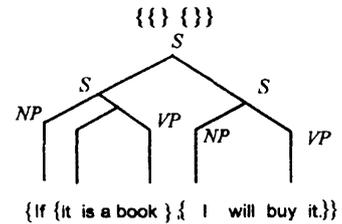
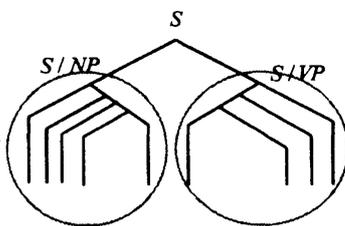


II-型文章

$clus\{\{\circ\}\{\circ\}\} = 2$ となる文章。

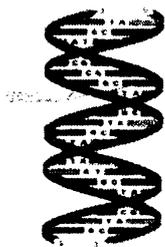
$\{\{\circ\}\{\circ\}\}, \{\{\{\circ\}\}\{\circ\}\}, \dots$

II-型文章の樹形図

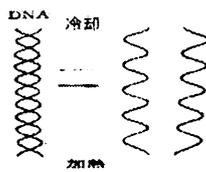


3. DNA モデル

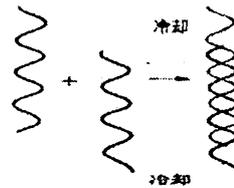
DNA は 4 種類の核酸 A, T, G, C の配列からなる 2 本の線の絡み合った螺旋である(下図左上)。基本構造は(1) $A \leftrightarrow T, G \leftrightarrow C$ となる相補性と(2) 突然変異である。相補性は下図右のように 2 種類のを考える。図の右側のずれをもつ相補性はトランスポゾン等を記述するとき重要となる。突然変異は欠損、挿入、置換等からなっている(下図左下)。Transition mutation とは互いに相補の位置にある核酸の入れ替えをいう(下図右下)([2])。



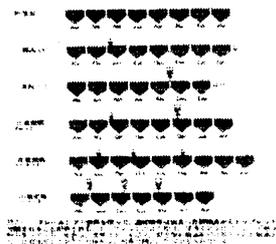
Normal denaturalization



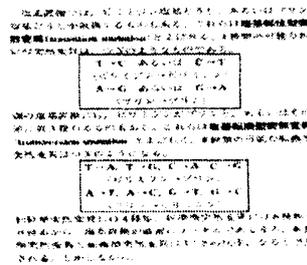
Anormal naturalization



Mutation in DNA



Transition mutation

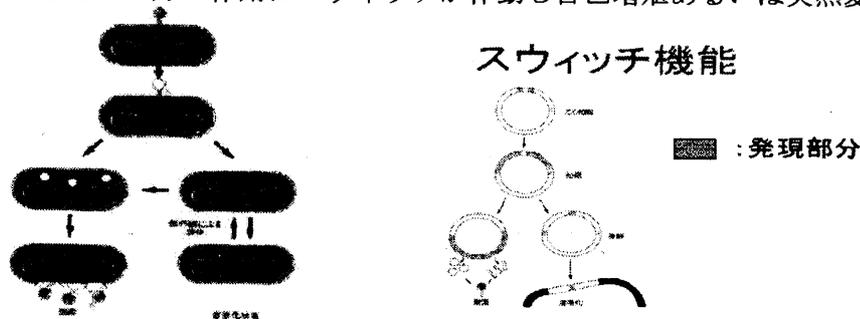


4. トランスポゾン

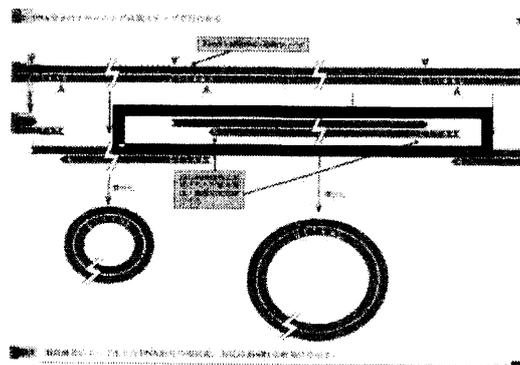
従来 DNA における進化は時間変化に注目した進化(これを垂直方向の進化とも言う)が中心に考えられてきたが、ミトコンドリアあるいはλファージのように自らの DNA を別の生物の細胞内に送り込むことにより自己増殖を行ったり、寄生された生物の DNA に突然変異をおこす現象が確認されており、このような DNA をトランスポゾン(動く遺伝子とも言う)と言っている。トランスポゾンによってなされる変異を従来の変異と区別して水平変異と呼んでいる(詳細は[2]を参照せよ)。



λファージの他生物細胞への侵入およびその突然変異の開始はつぎの手順で行われる。外的刺激に応じて一方の作用にスイッチが作動し自己増殖あるいは突然変異を生じる。



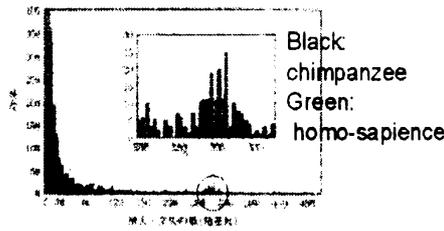
トランスポゾンはトウモロコシやショウジョウバエのような真核生物等の高等な生物にもみられる(下図)。DNA の一部を取り去りこれを円環にして運搬したもとの線状 DNA に戻してべつの DNA に進入することにより突然変異をおこす([2])。DNA コンピューティングにおけるスプライシングはこの作用に基づいて行われると思われる([3])。



スキームは次のようになる。上列、下列の核酸の配列およびその対応に注意する(下図左)。このような配列のおかげで円環と線分の両方に変化できる。



次ページの図はヒトとチンパンジーについて突然変異が生じている塩基の長さに対するその頻度数をグラフにしたものである。280 前後に小さなピークが確認されるがこれは ALU とよばれるトランスポゾンによる突然変異である。このことから知能にはトランスポゾンが深くかかわっていると考えられる([4])。

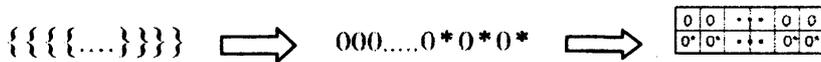


5. 整合括弧列文の DNA モデル

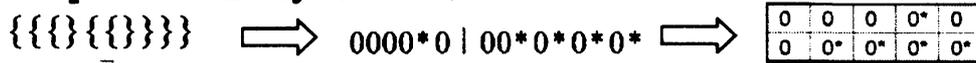
整合括弧列文のDNAモデルを構成する。現在、言語機能に対応するDNAの物質が何であるかは決定されるには到っておらず、我々の考察は暫定的なものと言わざるを得ない([5])。設定を述べる：

(1) **DNA の構成**：DNAを構成する要素は $0, 0^*$ であり、括弧の総数は偶数である。そこで順序をかえることなくその半分を上段に配列し残りの半分を下段に配列する：

Example 1(Primary sentence)



Example 2(Secondary sentence)



この配列法は多分に問題点を含んでいる。これが正当化されるかどうかは今後の研究に委ねられている。実際、言語の理解を信号伝達システムとして理解するときこの問題が決着されるものと思われる([8])。

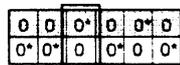
(2) **文章の分類**：文章の幾つかのクラスを述べる。引き続き開き括弧の列のあとに引き続き閉じ括弧の列が現れる文章を **primary sentence** といい(Example 1)、これ以外の文章を **secondary sentence** と言う(Example 2)。

(2) **T-operation**：Transposon から導かれる演算を次のように定め、これを **T-operation** (言語学では **Adjunction** ともいう)という。上列と下列をずらして貼り合わせる。共通部分でない部分の個数を **deg** という。下図では **deg** は 1 と 3 となる。



(4) **相補性**：相補性は $0 \Leftrightarrow 0^*$ により定める(下図)。この対応においては一般的な文章については相補性は成り立たない。それぞれの例を一つづつ与える。

Complementarities



$0 \Leftrightarrow 0^*$

complementary

Sentence without complementaritie



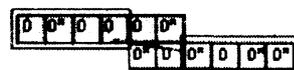
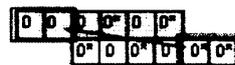
$\{\{\}\{\}\{\}\}$

この相補性のやぶれは **T-operation** により回復されることが示される。**T-operation** のある場合の相補性は斜線の同一視により定められる(下図)。

Transposon complementarities



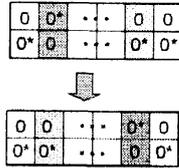
(1) Transposon of degree 1



(2) Transposon of degree 2,3

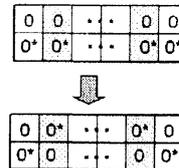
(5)幾つかの演算：次に基本的な演算を2つのべる：

(1) Substitution mutation



(2つの列を入れ替える)

(2) Transition mutation

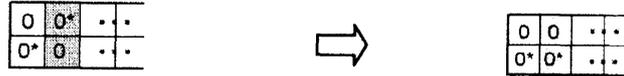


(相補の位置にある上下の元を入れ替える)

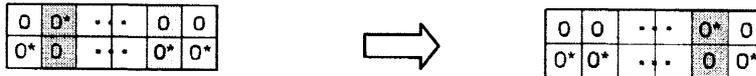
6. 整合括弧列文に対する変形規則

整合括弧列文に対する上記の演算は必ずしも整合括弧列文を整合括弧列文に移すとはかぎらない。そこで整合括弧列文を保存する演算を変形といい、これらの全体を変形規則という。演算が変形になるためには整合括弧列文の成立条件が満たされればよい。無条件で変形になるものをいくつか述べる。

(1) **Transition mutation deformation**：長文を幾つかの短文に分割したり、または幾つかの短文をまとめて長文にする作用がある。次の演算は変形である。



(2) **Substitution deformation**：長文を幾つかの短文に分割したり、あるいは短い文章をまとめて長文にする作用がある。次の演算は変形である。



7. トランスポゾン DNA モデル(I)(Static model)

T演算が文章生成においてどのような役割をはたしているかを考える。まず整合括弧列文章の生成規則を思い出す：

- (1)空集合に対して{ }を生成する。
- (2)文章 S に対して、{S}, {S}を生成する。

この生成規則から(1)はI型の文章を生成することを意味しており、(2)の2番目の生成はif文或いはbecause文のように複数の文章を結びつけることにより高度な内容の文章を作り出す方法を与えているといえる(次ページ図左)。このようにT演算は整合括弧列文のなかに多くみられる。実際に次の定理がなりたつ：

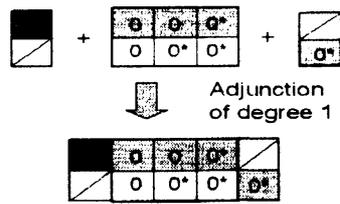
Minimal Construction Theorem(I)(Static model)

すべての整合括弧列文は次の演算により primary sentence に帰着される。

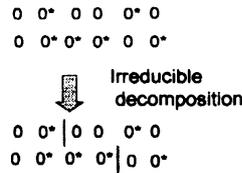
- (1)Irreducible decomposition into primary sentences
- (2)T-operation of degree 1

この内のいずれかひとつでもはずすと構成されない文章が存在する。この意味において minimal といえる。

ここで既約分解とは0と0*の個数を順番に数え始めて夫々の和が等しくなるときこれは一つの既約文章とみなし切り離す。この操作を何回か繰り返して短文の列に分解することを既約分解という(下図右)。



Irreducible decomposition



8. トランスポゾン DNA モデル(II)(Dynamical model)

つぎの定理がなりたつ：

Minimal Construction Theorem(II)(Dynamical model)

すべての整合括弧列文章は次の3つの演算を行うことにより primary sentence に帰着される：

- (1) Transition mutation,
- (2) Substitution,
- (3) T-operation

すべての演算は変形規則を作用することにより実現される。この内のいずれかひとつでもはずすと帰着されない文章が存在する。この意味において minimal といえる。

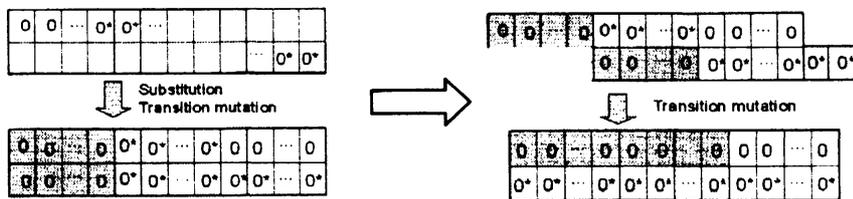
証明は難しくない。つぎの順序で primary sentence に帰される：

Step 1:最初に mutation と transition mutation を何回か行って下図左のように並べ替えることができる。

Step 2:つぎに T-operation 演算を行って下図右のように並べ替えることができる。

Step 3:最後に transition mutation を何回か行って primary sentence にできる。

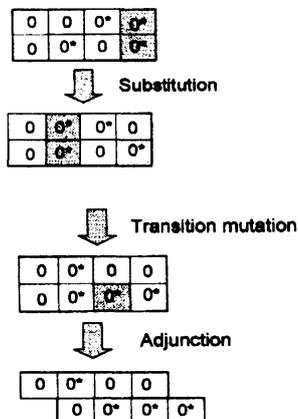
Reduction scheme



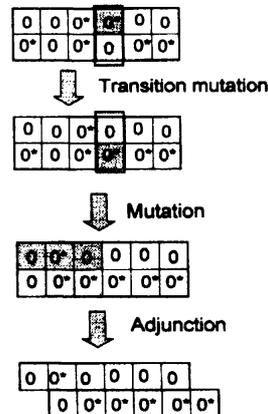
(同じ個数)

幾つか例を述べる。

Example (1)



Example (2)





最後に minimality condition にふれる。

- (1) 長さが3となる文章はすべて(1),(3)により primary sentence にできる
- (2) 長さが4となる文章は次の文章以外は(1),(3)により primary sentence にできる。

0	0	0*	0*
0	0*	0	0*

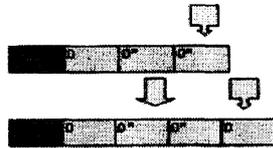
この文章に(2)を行うと primary sentence にできる。ここで文章の開き括弧の総数(これは閉じ括弧の総数でもある)を文章の長さという。

9. T-operation を用いた Turing machine の構成の試み

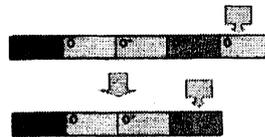
ここでは T-operation を用いて記述されるチューリング・マシンを構成する。チューリング・マシンの構成法は数多く知られている([1])。ここでは DNA の機能を基盤としてこれを構成する。DNA の2本のテープの元の並びは相補条件により束縛されている。この条件と T-operation を用いて一方の核酸列(これを上列という)にチューリング・マシンを構成する。もう一方の列は上列の補助の役割を果たしていると考え。基礎的なデータを述べる：

- (1) このテープには 0 と 0* の2種類のアルファベットが書き込まれるものとする。
- (2) ブランクも用意する(これは DNA ではイントロンの役割をはたす)。
- (3) 始点を \$ とかく。
- (4) 右シフトの存在を仮定する(下図左)。

(1) Right shift



(2) Left shift

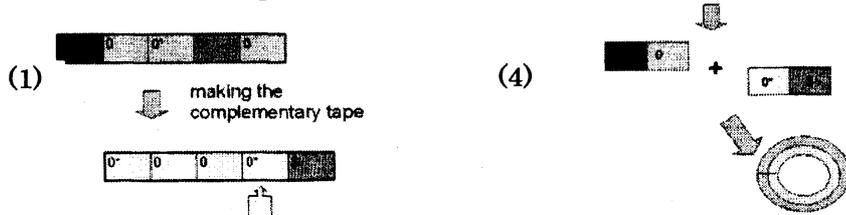


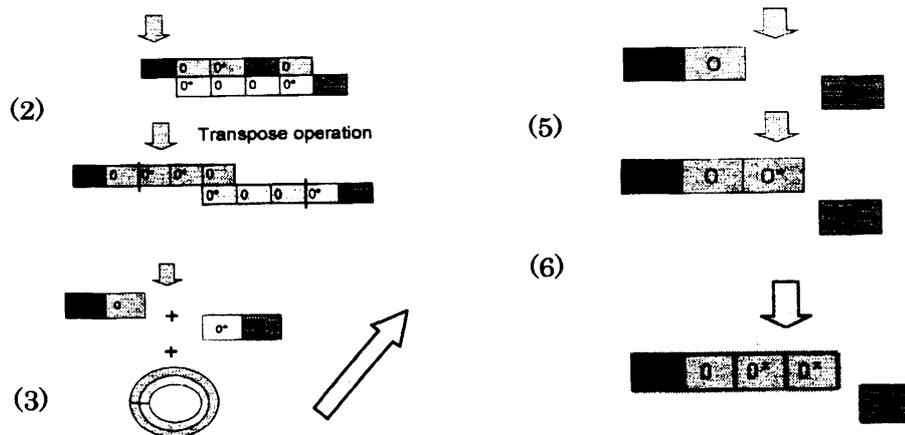
ここでは左シフトの存在を仮定しない(上図右)。以下これについて考える。基本的な疑問は次のものである：

「DNA ははたして左シフトを行い書き換えを行うであろうか？」

DNA はこの様に一度生成した核酸あるいは蛋白質を誤りとして遡って修正を行わないと考えるのが自然であろう。実際、これらの物質の個数は膨大であり遡った修正は徒に混乱を招くだけであろうと思われる。また、誤りという概念は存在しないからである。突然変異は誤りとも言えるからである。このために DNA はイントロンあるいは「ジャンク」をつくり、其の内からほんの僅かの有用な配列を選んでいるように考えられる。以下では T-operation がどのように左シフトを再現するかについて考える。

上記の演算を用いると上図左のようにテープに書き込まれる。上図右にある左シフトがどのように(1)相補性、(2)T-operation、(3)生成により実現されるかを見る。





このとき T-operation の作用の結果(3),(4)において「ジャンク」が作り出されていることに注意する。

10. Discussions

以上の考察から得られた結果をのべ、今後の課題等について述べる。

(1) 整合括弧列のつくる文章の生成規則およびチューリング・マシンの演算をトランスポゾンから導かれる T-operation を用いて記述することができた。トランスポゾンが知能の生成あるいは言語機能の生成と密接に関係することに注意すると、これが我々が日常何気なく行っている知的行為を理解する手法を提供するものと期待される。

(2) 次にチューリング・マシンにおけるトランスポゾンの役割について考える。ここでは DNA のもつ基本的な演算をもちいて新しいチューリング・マシンの提案を行った。この意味について少し述べる。DNA のもつ基本的な作用を基礎としてチューリング・マシンを構成することは極めて重要なことである。それは「人間をふくめた生物がどのように計算を行っているのか？」を理解することが出来るからである。我々のチューリング・マシンは非可逆なチューリング・マシンであり時間の非可逆性とも密接に関係している。人々の数理演算は個性のあるものと思われる。これを DNA レベルで考察することも可能である。また、人々はどのようにして計算間違いをするのか、これについて理解することも可能であろうと思われる。人々が数学的真理をどのように共有できるのかについてもここで構成されたチューリング・マシンと時間可逆なチューリング・マシンと比較することによりなされるであろう。これは Bennett の定理により保障されているといえる([6])。今後は論理構造の共有可能性等について考察することは興味あるものと思われる。

REFERENCE

- [1]米田正明、広瀬貞樹(他2名):オートマトン・言語理論の基礎近代科学社(2003)
- [2]ハートル、ジョーンズ:エッセンシャル遺伝学培風館(2003)
- [3]G.パウン、G.ローゼンバーグ、A.サローマ:DNA コンピューティングシュプリ
ンガー・フェアラーク東京(1999)
- [4]斉藤成也(他6名):遺伝子とゲノムの進化(シリーズ進化学2)岩波書店
- [5]S. Paabo: Humann evolution, Millennium issue, Science14~116(2000)
- [6]C.H.Bennett: Logical reversibility of computation, IBM.J.Res. Dev.,17,525-532(1973)
- [7]N.Chomsky:Context-free grammar and pushdown storge, Quarterly Prog. Rept.
No. 65,187-194, MIT Res. Lab.Elect, Cambridge, Mass(1962)
- [8]山本 雅(他1名):シグナル伝達・イラストマップ羊土社(2004)