

卒業研究

グラフ言語理論のパターン認識への応用 - 甲状腺画像の認識—

日本大学文理学部応用数学科
夜久研究室 齋藤有紀

パターン認識とは

種々の対象をあらかじめ想定されたパターンの
のいずれかを判定する処理
対象としては図形や文字・音声などがある

パターン認識の研究における二つのアプローチ

- ・ 解析学的アプローチ
- ・ 言語理論アプローチ

今回は言語理論的アプローチを取り扱い
甲状腺をモデルとして識別する

解析学的アプローチとは

言語理論でいう文法をカテゴリーごとに用意し、未知のパターンの特徴系列がこの文法に適合するか否かで、その属性を決定する方法

甲状腺画像のパターン認識

甲状腺の形状を認識し、正常か異常かを判別する。異常な甲状腺には形状にさまざまな変化があらわれる。

あらかじめ用意されている正常な甲状腺の特徴形状のパターンに適合するか否かで判別を行う。

グラフ・・・ネットワークの数学的モデル

言語理論・・・文法により文字や図形の集合を解析する理論

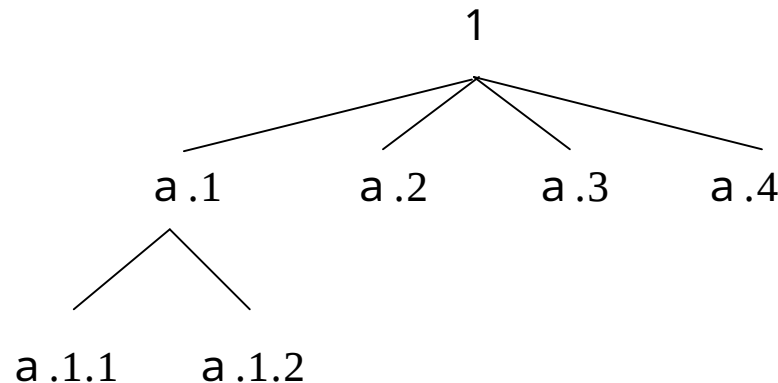
パターン認識・・・種々 図形・文字・音声をあらかじめ
想定されたパターンのいずれかを判定する

甲状腺・・・内分泌腺の1つ。のどぼとけの下方に位置し、気管の両側にある。甲状腺ホルモンを分泌して代謝を調節する。外見は赤みの強い桃色で蝶のような形をしている。

定義 2 . 1 木領域

木領域 D :

- (1) 1 D であり、1 は木の根である。
- (2) $a \in D$ のとき、 a に直接つづく 節点 (子) は $a.i$ の形式である。ここで i は自然数 N の要素である。
- (3) $a.i \in D$ のとき、 $a \in D$ かつ $a.1, a.2 \cdots a.i \in D$ である。
- (4) $a.i, a.j \in D$ のとき、節点 $a.j$ が節点 $a.i$ の右にあれば、 $i < j$ である。



定義 2 . 2 木

木 T は、2 字組 (D, L)

D : 木領域

$L : D \rightarrow \Sigma$ のラベルの関数

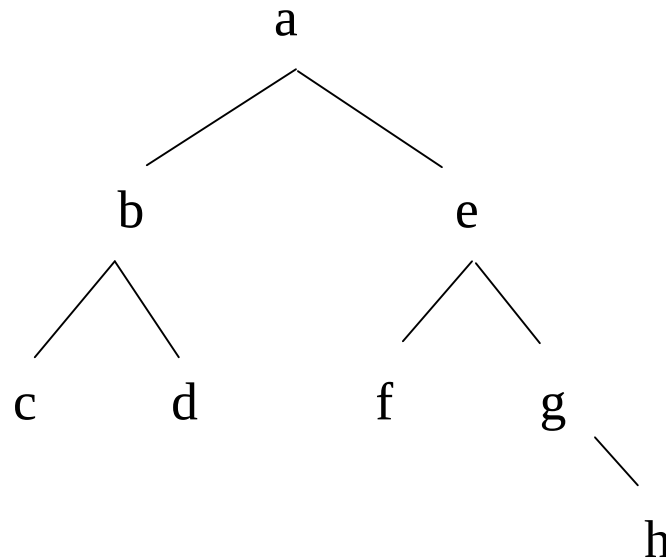
Σ : ラベルの集合

木Tの例

木 $T = (D, L)$ を次のそれぞれの式により定めると次のようになる。

$D = \{ 1, 1.1, 1.2, 1.1.1, 1.1.2, 1.2.1, 1.2.2, 1.2.2.1 \}$

$L = \{ (1,a), (1.1,b), (1.2,e), (1.1.1,c), (1.1.2,d), (1.2.1,f), (1.2.2,g), (1.2.2.1,h) \}$



定義 2 . 3 本文法

本文法 G_t は、4 字組 (V_N, V_T, P, S)

V_N : 非終端記号の有限集合

V_T : 終端記号の有限集合

P : 書き換え規則の有限集合

S : 開始木の有限集合

木 T_1 が G_t の規則 P により書き換えられて T_2

となるとき、 $T_1 \rightarrow T_2$ と書く。また $T \rightarrow T_1$

、 $T_1 \rightarrow T_2$ 、 $T_2 \rightarrow T_3$ 、 $\dots$ 、 $T_n \rightarrow T$ の

とき $T \rightarrow T$ と書く。

定義 2 . 4 木言語

木文法 G_t により生成される木言語 :

$$L(G_t) = \{ T \mid T = (D, L), \quad a \in D, L(a) \subseteq V_T \\ \text{かつ} \quad T_i \in S, T_i \in T \}$$

定義 2 . 5 正規木文法

木文法 G_t は以下の条件を満たす

- ・ 開始木の集合 S は、節点数 1 の木の有限集合である
- ・ アルファベット $V_N \cup V_T$ はランク付きのアルファベットであり、 $(V_N \cup V_T, r)$ と書ける
- ・ 規則 P は図 1 に示す形式であり、
 $Y, X_1, X_2, \dots, X_n \in V_N, X \in V_T, n \leq r(X)$

図 1 の規則を用いて木Tの葉 a を書き換える
 $T \rightarrow T'$, $L(a) = Y$,
 $D' = \{ D \quad a.i \quad | \quad i = 1, 2, \dots, n \}$
 また、 L' は $L'(a) = X$, $L'(a.i) = X_i$

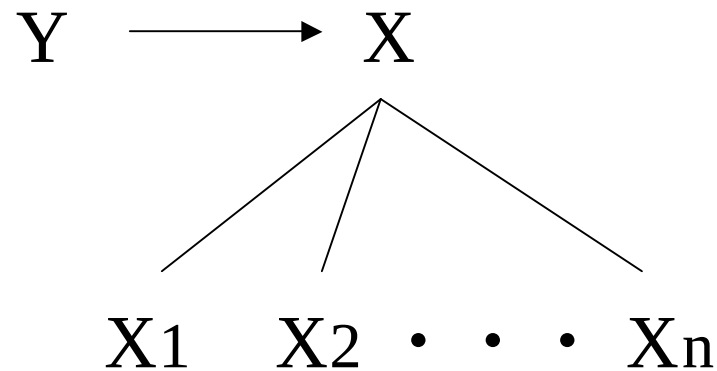
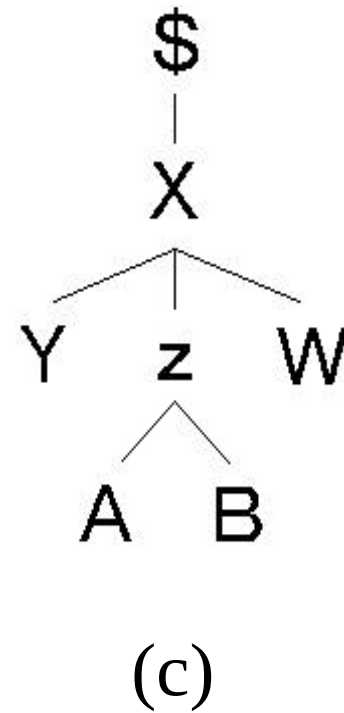
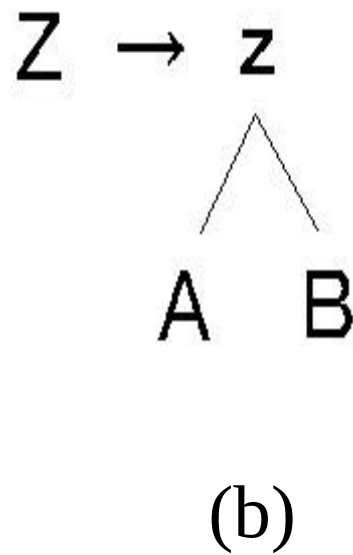
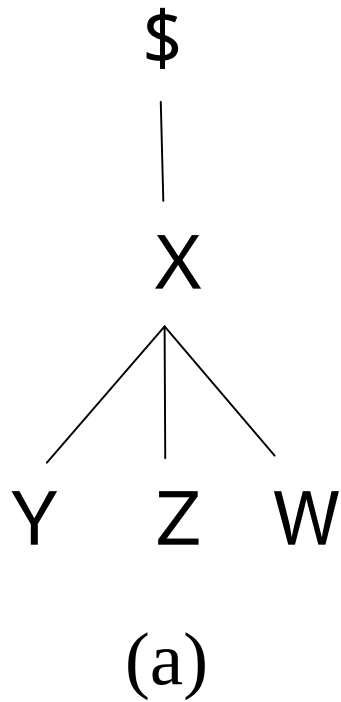


図 1 . 正規木文法の規則 P

正規木文法による書き換えの例



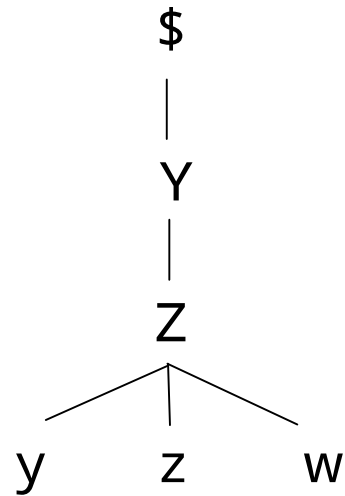
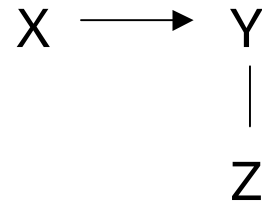
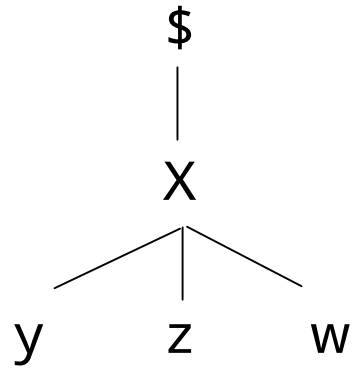
定義 2 . 6 自由木文法

$$(1) X \rightarrow YZ \quad , 1$$

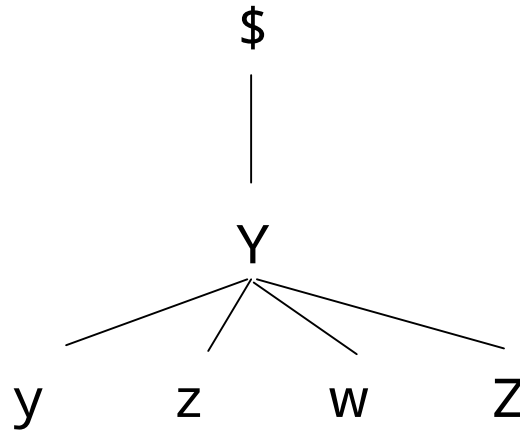
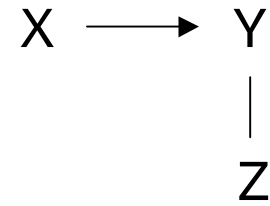
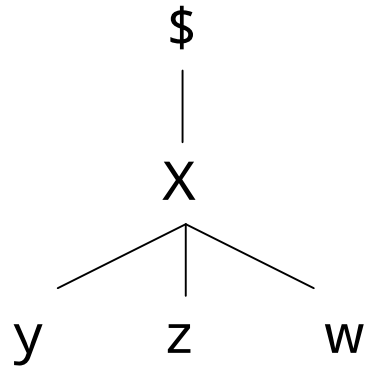
$$(2) X \rightarrow YZ \quad , 2$$

$$(3) X \rightarrow x \quad , 3$$

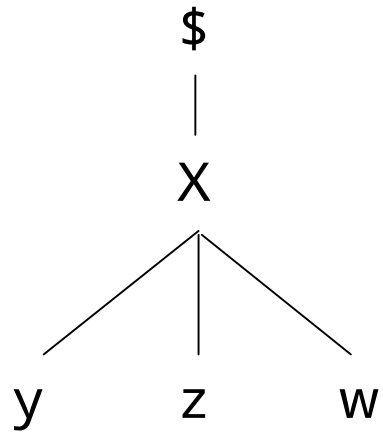
タイプ 1 ($X \rightarrow YZ$, 1)



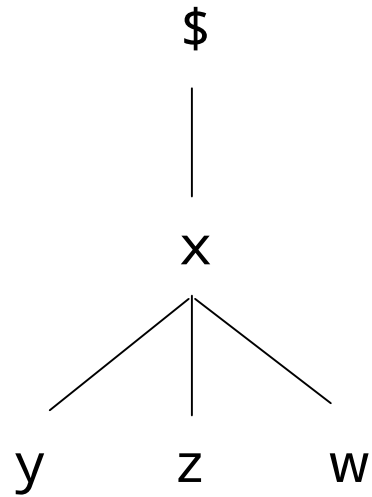
タイプ 2 ($X \rightarrow YZ$, 2)



タイプ 3 ($X \rightarrow x, 3$)



$X \longrightarrow x$



甲状腺画像認識処理

甲状腺の画像診断によって取り込んだ画像を識別することを目指す
20 × 20サイズに細分化した画像で、
2値画像を用いる

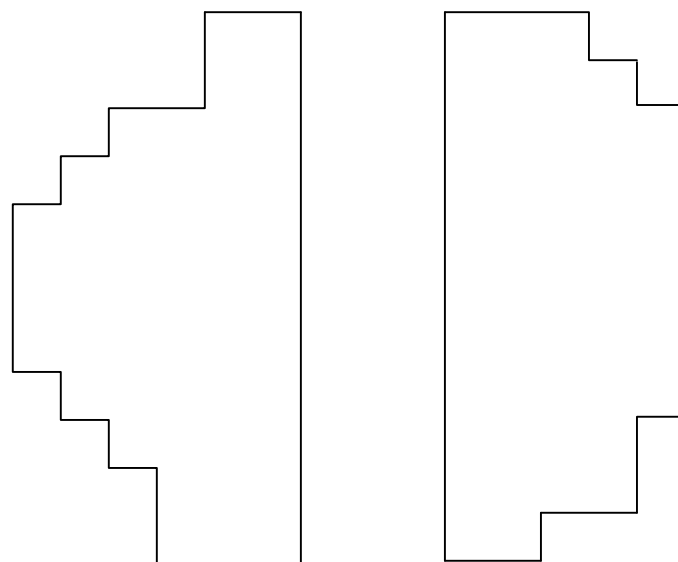
異常な甲状腺は形状に変化が現われる

凹凸、肥大、分裂、接触

がそれに当てはまる

次の手順に従って甲状腺の形状を
識別する

正常な甲状腺画像



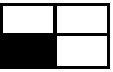
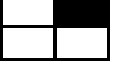
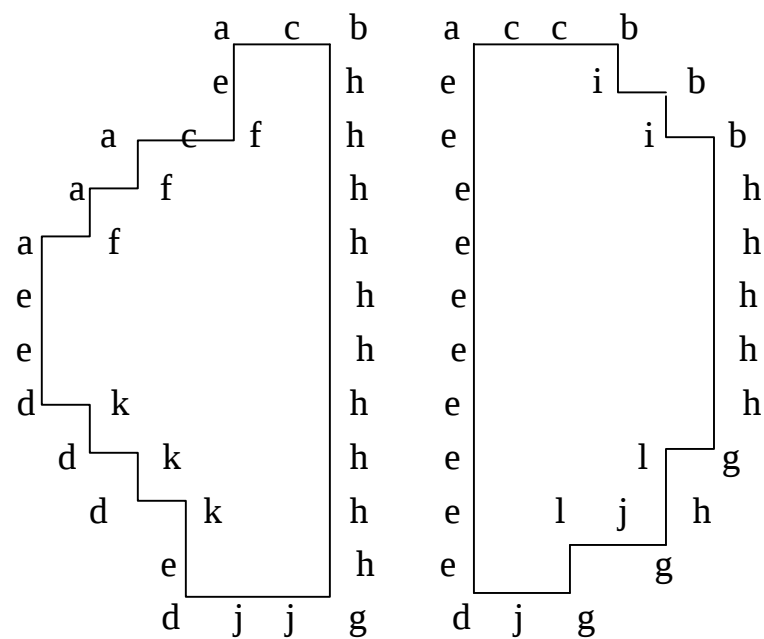
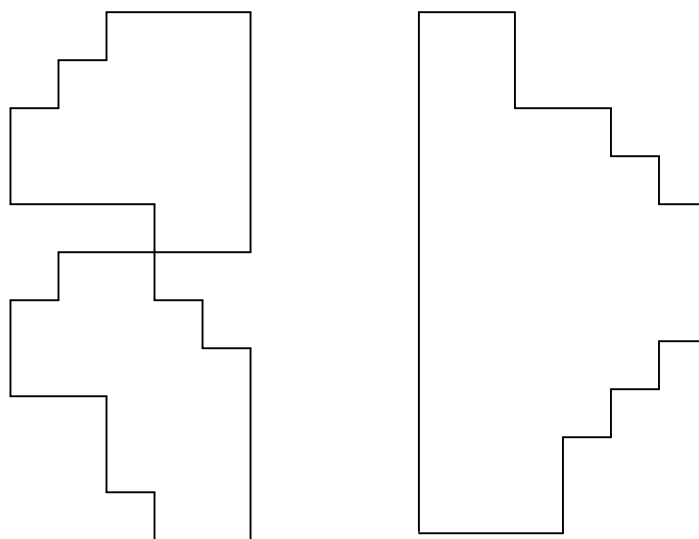
	境界の形状	形状コード		境界の形状	形状コード
1		a	8		g
2		b	9		*
3		c	10		h
4		d	11		i
5		e	12		j
6		*	13		k
7		f	14		l

表1．輪郭の形状と対応するコード

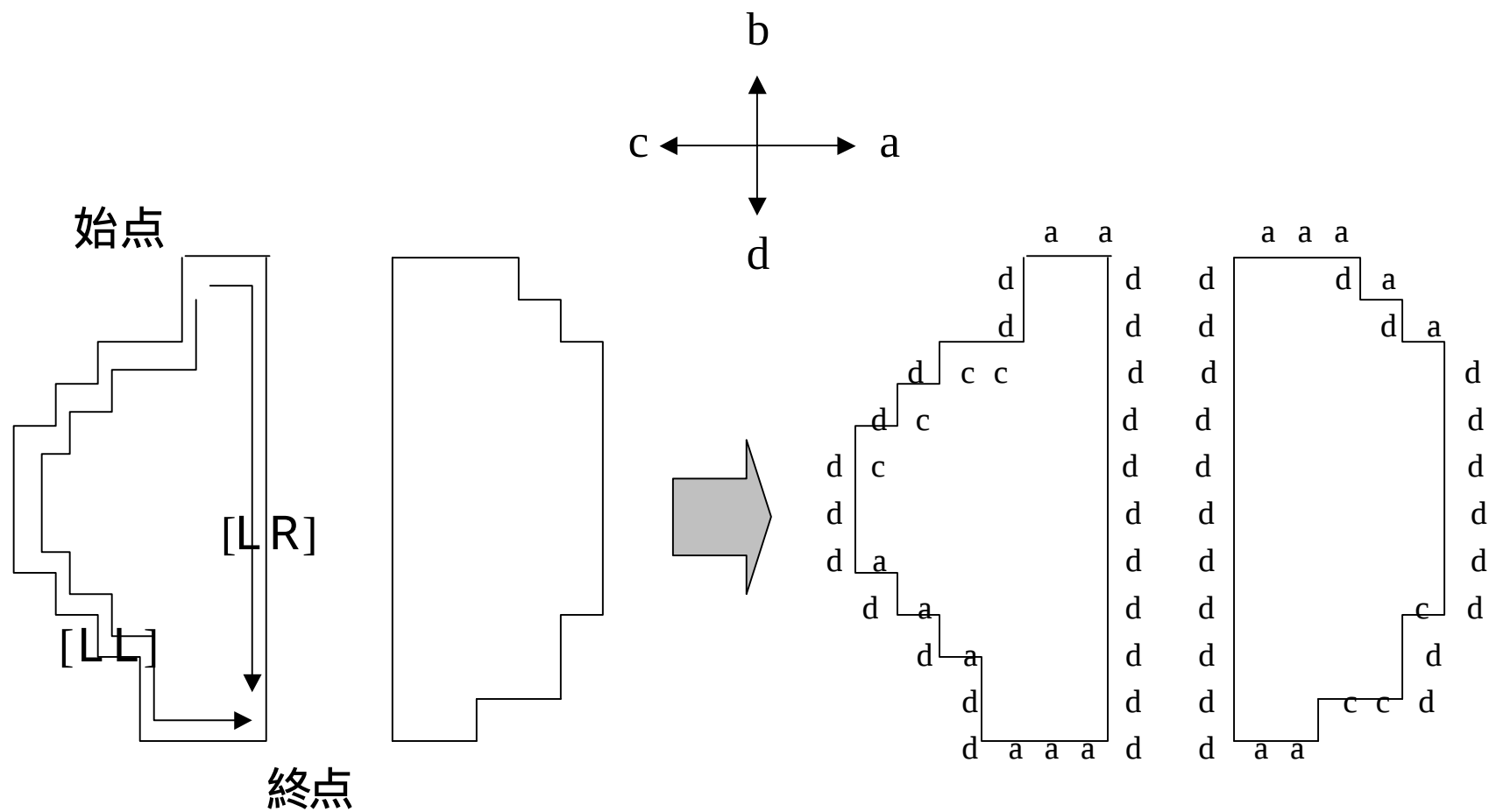
形状コード化された甲状腺画像



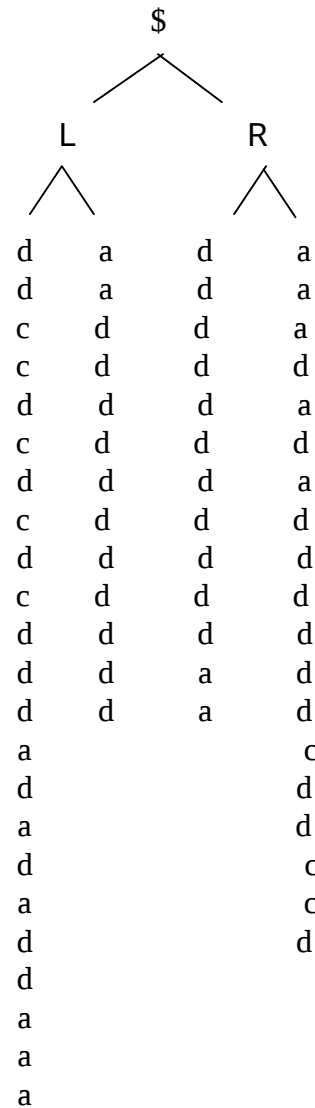
接触



チェーンコード

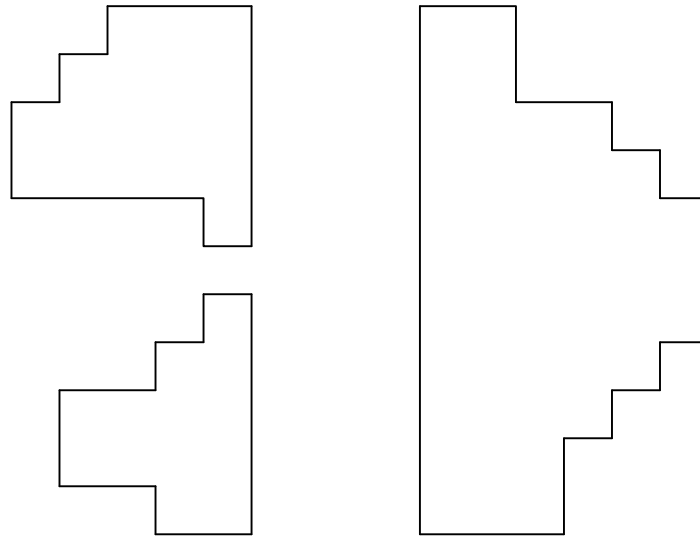


チェーンコードの木表現



[LL] [LR] [RL] [RR]

分裂

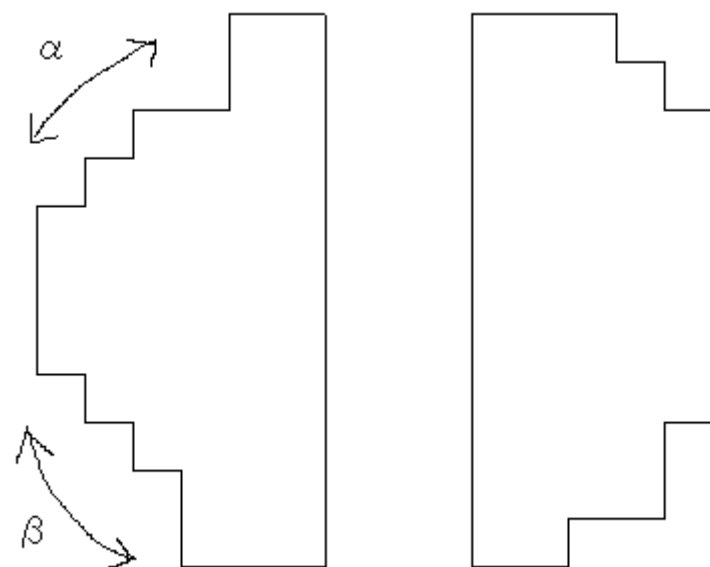


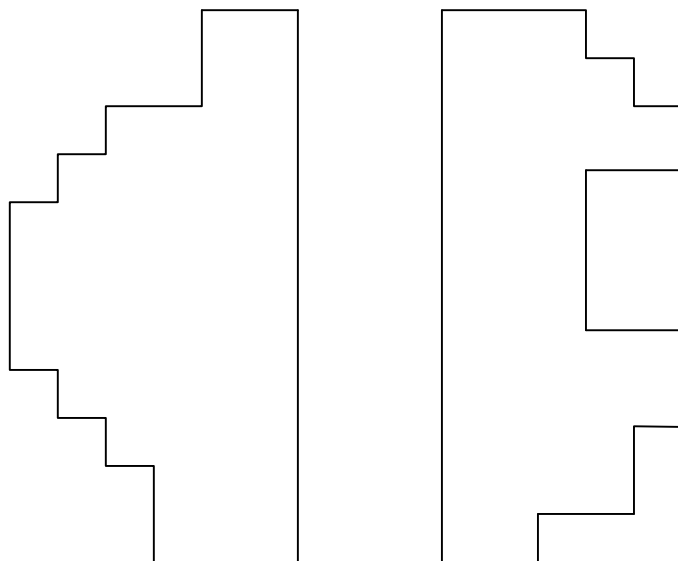
構文解析手続き

正常な形状の甲状腺の木表現を生成
する自由木文法

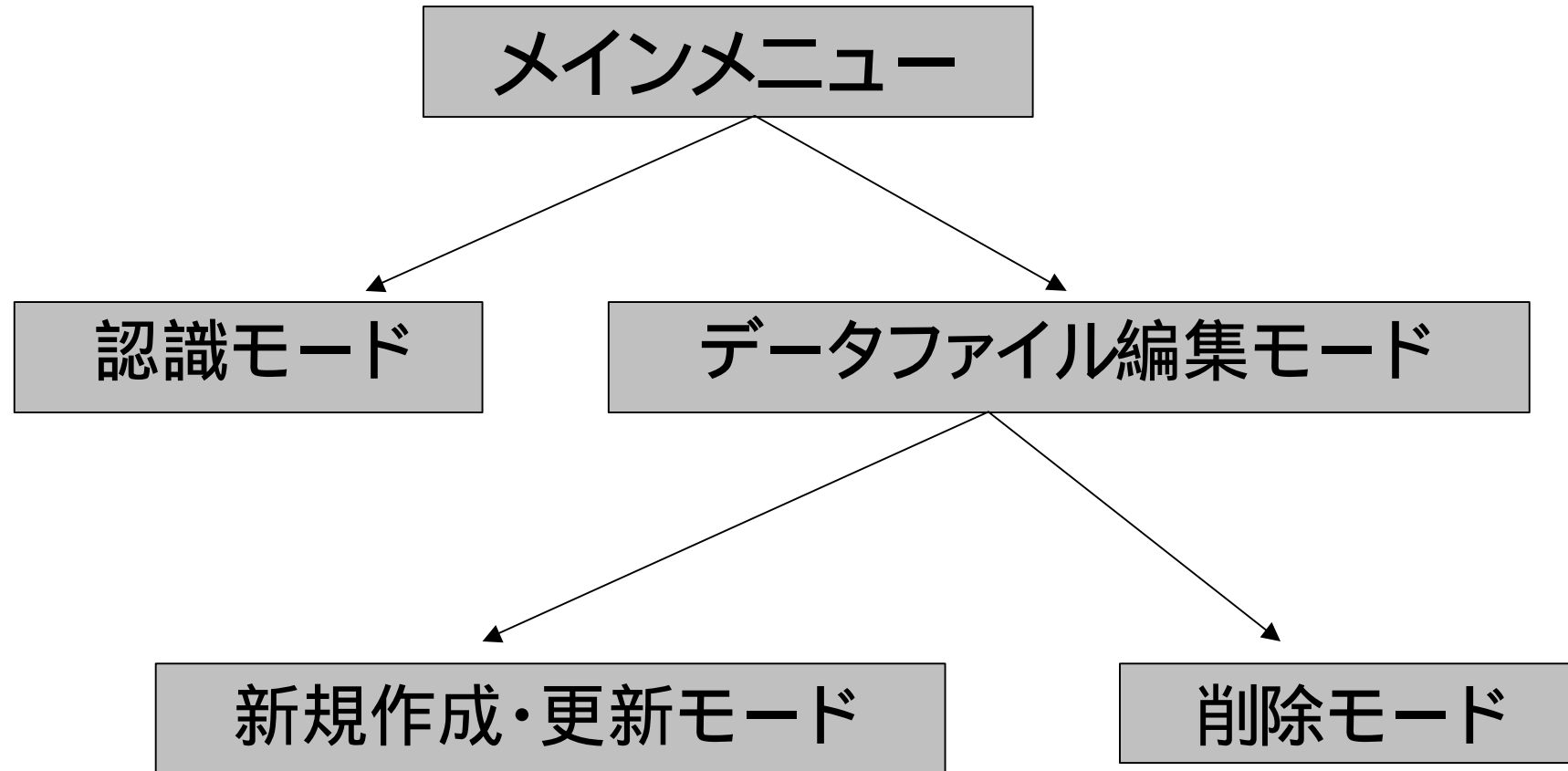
$$G_t = (V_N, V_T, P, S)$$

をもとに構文解析手続きを行う

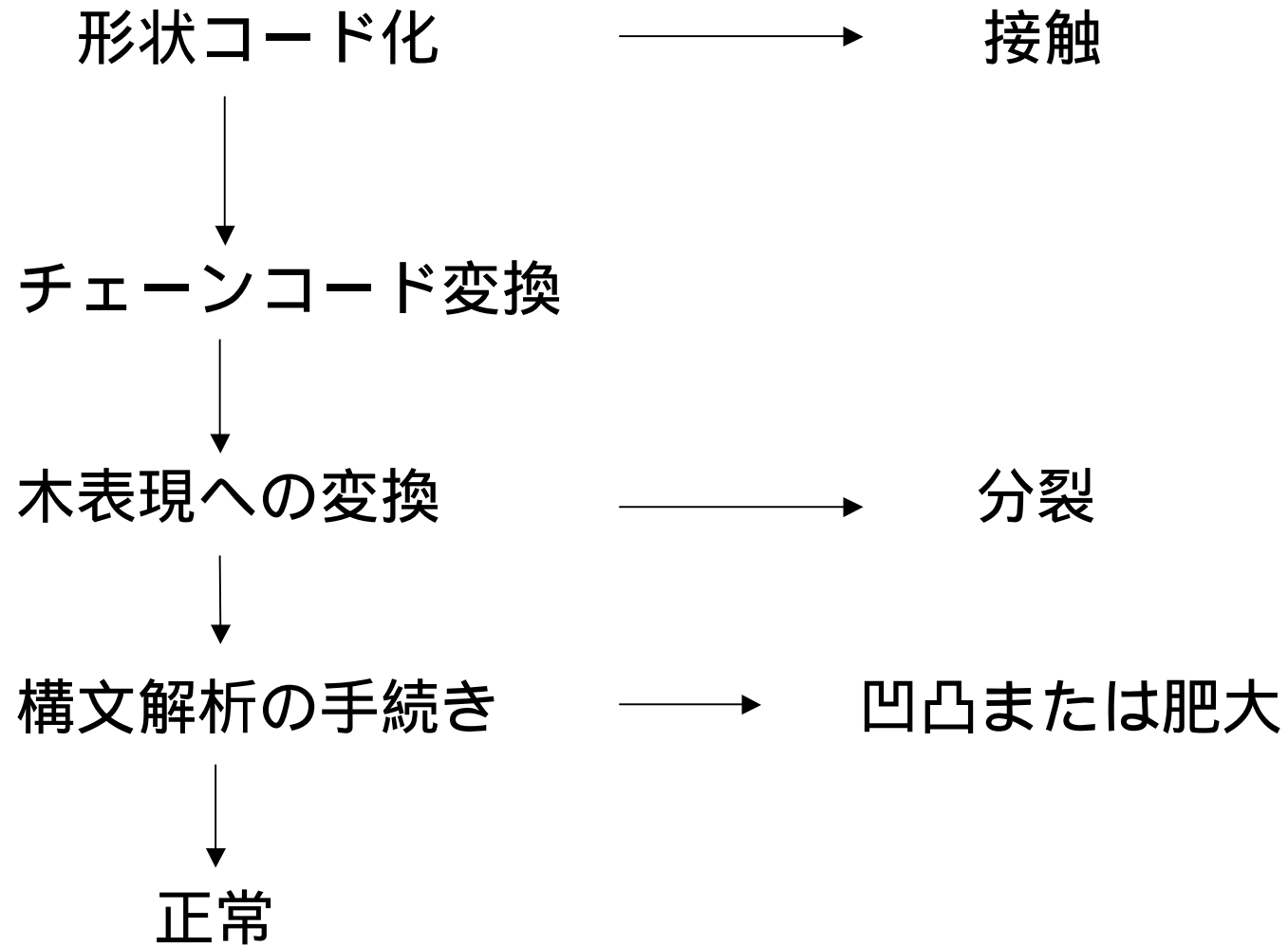




画像認識システム



認識モード



データファイル編集モード

新規作成モード

ファイルの新規作成、更新

データの編集

データのSave

削除モード

ファイルの削除

結論

このシステムでは、正常な形状をもつ甲状腺および異常な形状をもつ甲状腺いずれにおいても識別できるということが分かった。

おわりに

現在の段階では、正常な甲状腺と接触や分裂などの異常な甲状腺とを識別するのみにとどまっている。

今後の課題としては、その後の引き続きの適切な処理、画素数の増加、システムの信頼性を高めるためのより多くの甲状腺シンチグラムによるテストを行うことである。

参考文献

1. 添田明宏, 甲状腺画像認識システム, F A R S Tソフトウェア報告書 (1985)
2. 細田英和, 「グラフ言語理論のパターン認識への応用」, 日本大学文理学部応用数学科卒業研究報告書 (1998)
3. 南昌樹, 「グラフ言語理論のパターン認識への応用」, 日本大学文理学部応用数学科卒業研究報告書 (1999)
4. 河西朝雄著, J A V A入門, 技術評論社
5. 基礎から学ぶJava言語プログラミング, 電子開発学園出版社
6. 丸の内とら, 初体験Java, 技術評論社
7. 結城浩, Java言語プログラミングレッスン上, SOFTBANK
8. 結城浩, Java言語プログラミングレッスン下, SOFTBANK
9. はじめて学ぶP A S C A L, 共立出版株式会社
10. 真野芳久, P a s c a lプログラミングの基礎, サイエンス社
11. 和賀井敏夫, 佐久間浩, 乳房・甲状腺アトラス, ベクトル・コア