

卒業論文

韻律情報を用いた
句境界検出に関する研究

東北大学 工学部 電気・情報系
丸岡(東) 研究室 4年

粟津辰功

目次

1 序論	2
1.1 研究の目的と背景	2
1.2 複数の句境界候補を生成する原理	2
1.3 本論文の構成	2
2 GA 法のアプローチ	5
2.1 実際の生物	5
2.2 数学的モデル化	6
2.2.1 Step 0 遺伝子の設定、初期集団の発生	7
2.2.2 Step 1 各遺伝子の環境への適応度の評価	7
2.2.3 Step 2 評価値の評価に基づき淘汰	7
2.2.4 Step 3 遺伝子の増殖	8
2.2.5 Step 4 遺伝子の交鎖	8
2.2.6 Step 5 突然変異	8
2.2.7 Step 6 Step 1 に戻る	9
2.3 GA 法の特徴	9
3 句境界検出のための GA 法のアプローチ	10
3.1 アプローチ	10
3.1.1 Step 0 遺伝子の設定、初期集団の発生	10
3.1.2 Step 1 各遺伝子の示す句境界と、実際のピッチパ タンとの相似度の評価	11
3.1.3 Step 2 前段の評価値に基づき淘汰	11
3.1.4 Step 3 遺伝子の増殖	12
3.1.5 Step 4 遺伝子の交鎖	12
3.1.6 Step 5 突然変異	12
3.2 計算量	12
4 GA 法を用いた句境界検出実験	14
4.1 実験環境	14
4.2 遺伝子数と世代と収束性の関係	14
4.2.1 実験方法	14
4.2.2 結果の評価法	14
4.2.3 結果・考察	15

4.3	Open data と Closed data の句境界検出実験	15
4.3.1	実験方法	15
4.3.2	結果・考察	15
4.4	100 位候補までとったときの句境界検出率	15
4.4.1	実験方法	23
4.4.2	結果・考察	23
5	結論	27
5.1	結論	27
5.2	今後の課題	27
	謝辞	28
	参考文献	29

第3章では、本研究で使用する句境界検出のアルゴリズムを述べる。
第4章では、句境界検出の実験の方法と結果について述べる。
第5章で、本研究の結論について述べ、今後の課題について考える。

2 GA 法のアルゴリズム

GA 法は、遺伝的アルゴリズムの名の示す通り、実際の生物の進化や遺伝をモデル化した最適化法で、特に組み合わせ問題に有効な手法である。

2.1 実際の生物

実際の生物は図 2 に示すようなサイクルで生きている。

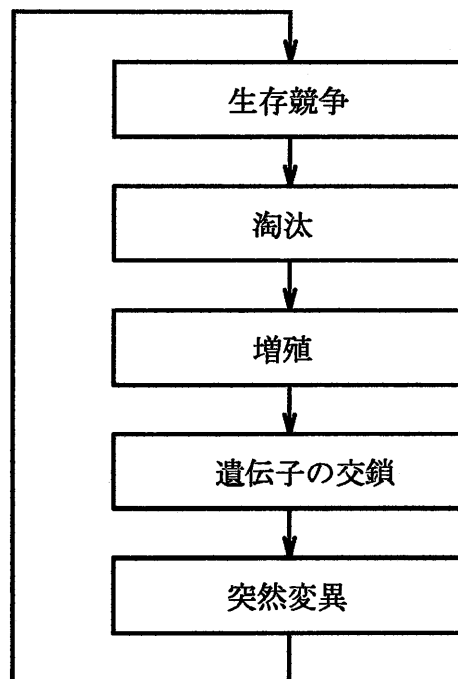


図 2: 実際の生物のサイクル

- Step 1 生存競争
- Step 2 淘汰
- Step 3 増殖
- Step 4 遺伝子の組み替え
- Step 5 突然変異

このサイクルにより、種としてのパフォーマンスが向上するのである。

2.2 数学的モデル化

これを数学的にモデル化したものがGA法で、図3のようになる。

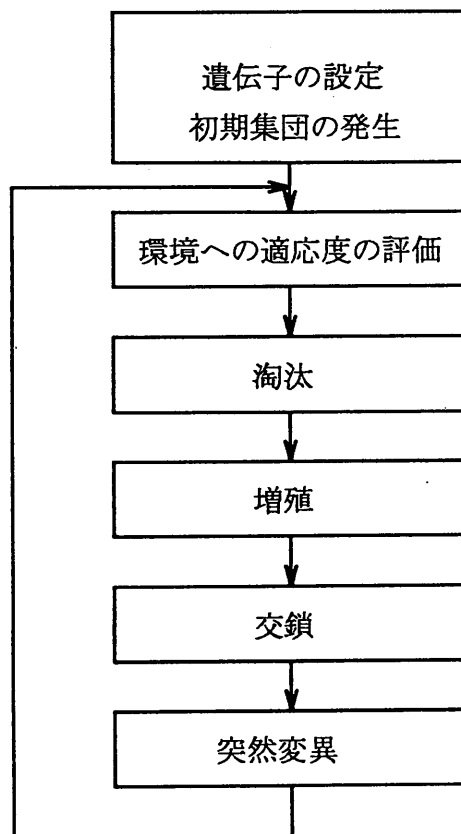


図 3: GA 法のアルゴリズム

- Step 0 遺伝子の設定、初期集団の発生
- Step 1 各遺伝子の環境への適応度の評価
- Step 2 Step 1 の評価に基づき淘汰
- Step 3 遺伝子の増殖
- Step 4 遺伝子の交鎖
- Step 5 突然変異

Step 6 Step 1 に戻る

このサイクルを世代と呼ぶ。

次に、各 Step について詳しく説明する。

2.2.1 Step 0 遺伝子の設定、初期集団の発生

GA 法では、遺伝子と呼ばれる記号列を用いて最適化を行う。まず、対象の事象をどのように遺伝子に記号列化するか設定し、そのコーディングの規則に従い、いろいろな個性を持った遺伝子の集合を発生させる。

各遺伝子は、基本的には同じ遺伝子型を持つが、細部の違いにより多数の個性をもつ遺伝子が現れる。この遺伝子という内部表現が、組み合わせ問題をコーディングしやすいので組み合わせ問題に適したアルゴリズムと考えられる。

2.2.2 Step 1 各遺伝子の環境への適応度の評価

各遺伝子が与えられた環境に、どれだけ適応してるか評価する。この評価を行う評価関数は、単なる記号列に過ぎない遺伝子と実際の事象とを結びつけるもので、GA 法で最も重要な役割を果たす。

2.2.3 Step 2 評価値の評価に基づく淘汰

Step 1 の評価に従い、次世代に生き残れる遺伝子を選ぶ。すなわち、環境に適応していない遺伝子を除去する。(図 4)

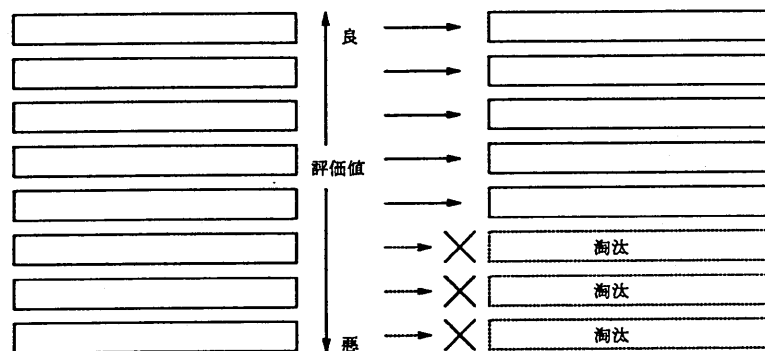


図 4: 淘汰

2.2.4 Step 3 遺伝子の増殖

Step 2 により、遺伝子数が減少するので、それを補うために遺伝子を増殖させる。(図 5)

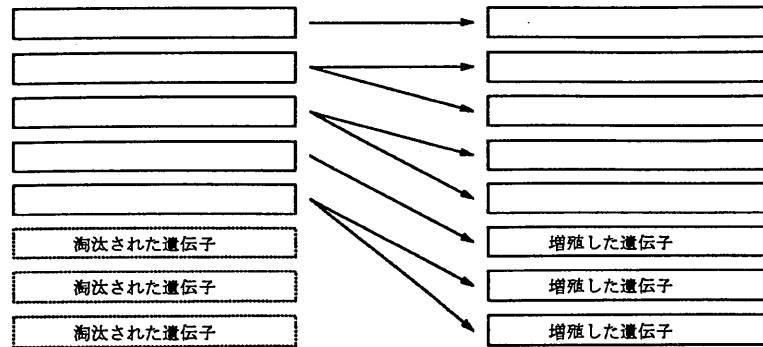


図 5: 増殖

以上の操作のみでは、遺伝子の初期集団の最適値以上の評価を持つ遺伝子は得られない。そのため、次の2つの操作が重要な意味を持つ。

2.2.5 Step 4 遺伝子の交鎖

遺伝子集団から特定の遺伝子対を選び、特定部位を入れ換える。(図 6) これにより、部分的に良い評価をもつ遺伝子から、より良い評価をもつ遺伝子が生まれる可能性がでてくる。

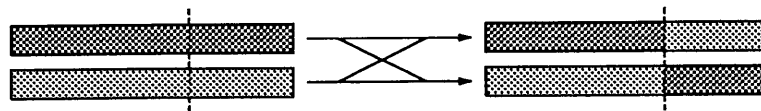


図 6: 遺伝子の交鎖

2.2.6 Step 5 突然変異

遺伝子の記号列の特定部位を変化させたり、記号列をずらしたりする。(図 7) これにより、突発的に良い評価をもつ遺伝子が生まれる可能性がある。

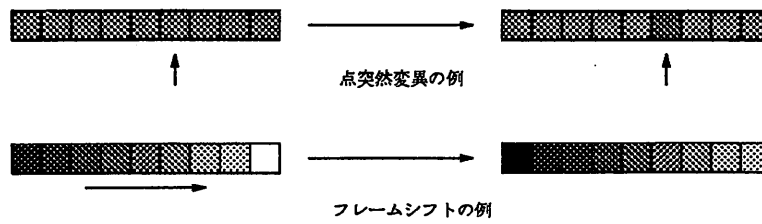


図 7: 突然変異

2.2.7 Step 6 Step 1 に戻る

以上のサイクルを、ある回数行う。そして、最後に残った遺伝子を近似解と見なす。

2.3 GA 法の特徴

- 複数の遺伝子が、相互作用しながら最適解に収束するため、収束が速く、これはSA 法など、単一の点で最適化を行うアルゴリズムを何回か行うこととは本質的に異なる。

3 句境界検出のための GA 法のアルゴリズム

3.1 アルゴリズム

以下に本研究で使用した、句境界検出のための GA 法のアルゴリズムを示す。(図 8)

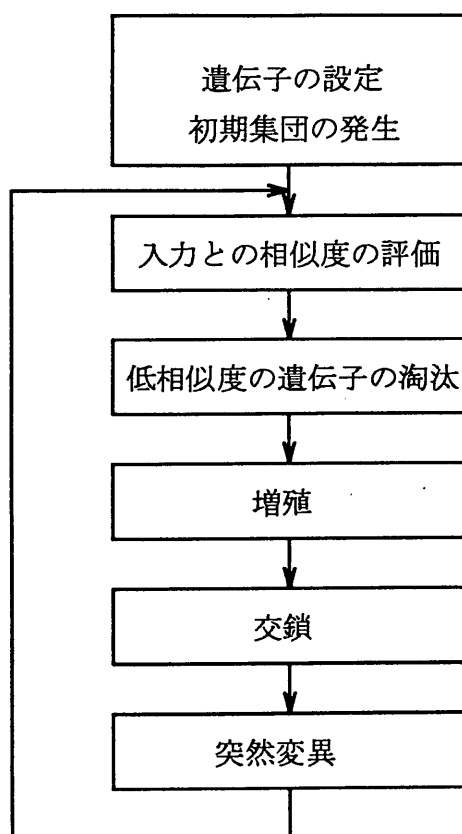


図 8: 句境界検出のための GA 法

3.1.1 Step 0 遺伝子の設定、初期集団の発生

遺伝子を、入力時系列長の配列で表わし、句境界を正の数で、それ以外を-1で表わすとする。(図 9) このコーディングの規則に従って、初期集団を生成する。

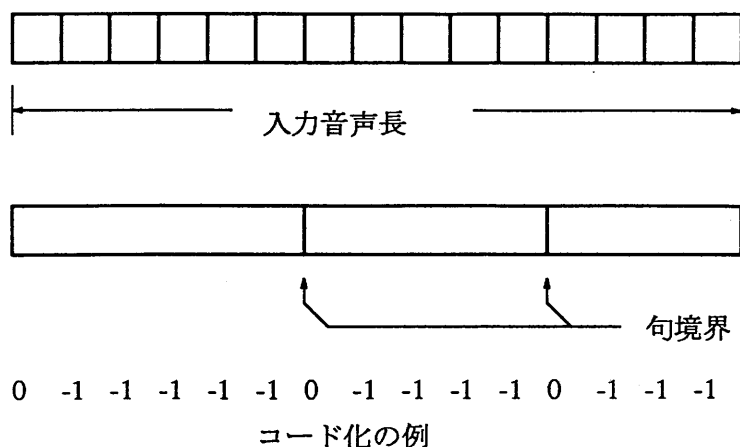


図 9: 遺伝子の設定

3.1.2 Step 1 各遺伝子の示す句境界と、実際のピッチパターンとの相似度の評価

遺伝子の示す句境界と、実際のピッチパターンとの相似度の評価関数として DP 距離を定義する。これは、One-Stage DP 法で用いられる評価関数を拡張したもので、遺伝子によって句境界が指定されたとき、入力音声のピッチパターンとテンプレートピッチパターンとを、句ごとに最小 2 乗誤差基準で DP マッチングしたときの 2 乗誤差の和である。この DP 距離が小さいほど、遺伝子の示す句境界と実際のピッチパターンとが似ていると考えられる。

また、ふつう、実際の音声は時間的に非線形に伸縮されて発声されるが、伸縮が極端に長すぎたり、短すぎたりすることは考えにくいので、DP マッチングの時間的伸縮に制限を設ける。本研究での傾斜制限は $[1/2 \sim 2]$ で、これにより各句は、テンプレートピッチパターンの長さの $1/2$ 倍から 2 倍までの時間的な非線形伸縮を許されることになる。

3.1.3 Step 2 前段の評価値に基づき淘汰

Step 1 の評価値をもとに、劣悪遺伝子、すなわち実際のピッチパターンにそぐわない句境界を示している遺伝子を除去する。評価値が小さいほど良い評価なので、評価値の逆数に比例する確率で遺伝子を次世代に残

すとする。

3.1.4 Step 3 遺伝子の増殖

Step 2 で減少した遺伝子を補充する。評価値の小さい遺伝子を優先的に増殖させるようにする。

3.1.5 Step 4 遺伝子の交鎖

遺伝子集団のうちの何割かの遺伝子を交鎖させる。交鎖させる遺伝子と交鎖箇所は乱数で決定する。

3.1.6 Step 5 突然変異

毎世代、幾らかの遺伝子に突然変異をおこす。

遺伝子の記号列の1点のみが変化する、点突然変異では時間伸縮の傾斜制限のため、実際にはありえないような句境界を示す致死遺伝子を作りやすいので、遺伝子の記号列が左右に幾つかずらず、フレームシフト突然変異を用いる。

以上のサイクルを、あらかじめ決められた世代だけ繰り返す。

このサイクルの中で作られる遺伝子をすべて調べて、評価値の良い順に、かつ独立な遺伝子を、句境界候補とする。

3.2 計算量

計算量は

$O(\text{遺伝子数} \times \text{世代} \times \text{入力音声長} \times \text{テンプレートピッチパターン数})$
である。

```

For i = start to end
  For k = 0 to number of templates
    For j = 0 to Jk
      
$$D(i, j, k) = \text{dist}(i, j, k) + \min \begin{cases} D(i-1, j, k) \\ D(i-1, j-1, k) \\ D(i-1, j-2, k) \end{cases}$$

    Next j
  Next k
Next i
dist(i, j, k) = (距離尺度関数)
DP 距離 =  $\sum_{\text{allsegment}} \min_k D(\text{end}, Jk, k)$ 

```

図 10: DP 距離の定義

4 GA 法を用いた句境界検出実験

4.1 実験環境

実験に用いた音声資料は、男性話者 1 名による自由発声連続音声 500 文章で、アクセント句を 4505 含む。その連続音声からの、ピッチ抽出には lag-window 法 (自動抽出) を用いる。上記 500 文章中 400 文章 (アクセント句 3601) のピッチ周波数と差分ピッチ周波数との 1 次結合を特徴量としてベクトル量子化を行い、得られた 2 つのパタンを、テンプレートピッチパタンとする。

4.2 遺伝子数と世代と収束性の関係

GA 法にはパラメーターとして、遺伝子数と世代の 2 つがある。この 2 つのパラメーターを変化させたときの句境界検出率と誤検出率を調べる。また、そのさい遺伝子数による句境界候補全体の評価も調べる。

4.2.1 実験方法

入力データとして、50 文章 (アクセント句 326) を用いて、遺伝子数 10/20/50、世代 10/20/50 の組み合わせ 9 通りについて、句境界検出実験を行う。

4.2.2 結果の評価法

句境界検出率と誤検出率は、視察によって与えられた句境界の $\pm 100\text{msec}$ 以内に検出された句境界が含まれる場合を正解として、次式で算出する。

$$\text{句境界認識率} = \frac{\text{検出された句境界の正解の数}}{\text{視察による句境界の総数}}$$

$$\text{誤検出率} = \frac{\text{検出された句境界の不正解の数}}{\text{検出された句境界の総数}}$$

これを、それぞれの候補につき計算し、もっとも良い句境界検出率、候補全体の句境界検出率の平均、もっとも悪い句境界検出率の 3 つで評価を行う。

4.2.3 結果・考察

上の句境界検出実験の結果を表 1 に示す。遺伝子数と世代と句境界検出率との関係では、遺伝子数と世代が多いほど検出率が良くなり、誤検出が減るのが分る。計算量に比例する遺伝子数×世代がほぼ同じである、遺伝子数 10 本－50 世代、遺伝子数 20 本－20 世代、遺伝子数 50 本×10 世代、の 3 つのデータに注目したときの結果を図 11、12 に示す。句境界検出率の向上は、遺伝子数にあまり関係しないが、遺伝子数が多いほど、誤認識が少ないことがわかる。これは、淘汰と増殖の操作で、遺伝子数が多いほど種々の遺伝子を次世代に残せる可能性が多いことによる影響であると考えられる。次に句境界候補全体の評価を考える。(図 13) 遺伝子数が少なすぎると、ローカルミニマムに陥り最適解に到達出来ないことがわかる。これも上記の理由と同様と思われる。また、遺伝子数に関わらず収束は非常に速く、10 世代程度で収束してしまい、それ以後の反復は無駄であることもわかる。以上より、世代を増やすより、遺伝子数を増やした方が良好な結果が得られることが分かる。

4.3 Open data と Closed data の句境界検出実験

テンプレート作成に使用した 400 文章 (Closed data) と評価用の 100 文章 (Open data) の句境界検出率を調べる。

4.3.1 実験方法

テンプレート作成に使用した 400 文章 (losed data) と評価用の 100 文章 (Open data) の句境界検出率を、遺伝子数 50 本－10 世代の GA 法で調べる。上の実験結果より遺伝子数 50 本－10 世代の組み合わせを選んだ。

4.3.2 結果・考察

結果を表 2 に示す。Open data / Closed data とも、One-Stage DP 法を上回る句境界検出率をしめす。まだ、テンプレート作成に使用していない Open data の方が良い結果をしめすことより、実験に使用したテンプレートピッチパターンが、アクセント句を良く代表していると考えられる。

4.4 100 位候補までとったときの句境界検出率

候補を多めにとったときの句境界検出率を調べる。

句境界検出率

	第3位候補			第10位候補		
	10世代	20世代	50世代	10世代	20世代	50世代
遺伝子数 10本	88.11%	87.66%	87.23%	89.23%	88.56%	89.23%
	77.72%	80.64%	79.14%	65.94%	72.44%	72.28%
	67.49%	73.31%	71.52%	44.17%	55.60%	55.38%
遺伝子数 20本	87.44%	87.89%	87.89%	89.01%	89.46%	89.68%
	83.18%	84.00%	85.42%	77.66%	79.55%	81.74%
	77.80%	80.26%	82.73%	63.45%	67.26%	71.30%
遺伝子数 50本	87.66%	87.44%	87.22%	89.23%	88.78%	88.56%
	85.95%	86.47%	86.17%	83.67%	84.21%	84.89%
	84.30%	85.20%	84.75%	74.43%	76.68%	79.59%
One-Stage DP法	86.99%					

誤検出率

	第3位候補			第10位候補		
	10世代	20世代	50世代	10世代	20世代	50世代
遺伝子数 10本	22.42%	20.85%	22.42%	23.54%	18.34%	20.63%
	31.01%	26.90%	30.26%	38.45%	32.62%	36.34%
	38.34%	31.84%	43.95%	51.79%	43.94%	48.20%
遺伝子数 20本	20.40%	21.15%	20.40%	17.93%	17.48%	15.92%
	26.00%	25.11%	24.51%	29.88%	28.07%	27.06%
	31.39%	28.92%	28.25%	42.60%	41.03%	37.89%
遺伝子数 50本	21.74%	22.42%	23.09%	16.14%	17.71%	18.38%
	24.58%	24.21%	24.66%	25.51%	25.15%	25.29%
	26.90%	26.45%	26.45%	35.65%	33.40%	31.83%
One-Stage DP法	21.77%					

最良の候補のみの句境界検出率/誤検出率 候補全体の平均の句境界検出率/誤検出率 最悪の候補のみの句境界検出率/誤検出率

表 1: 遺伝子数、世代と句境界検出率との関係

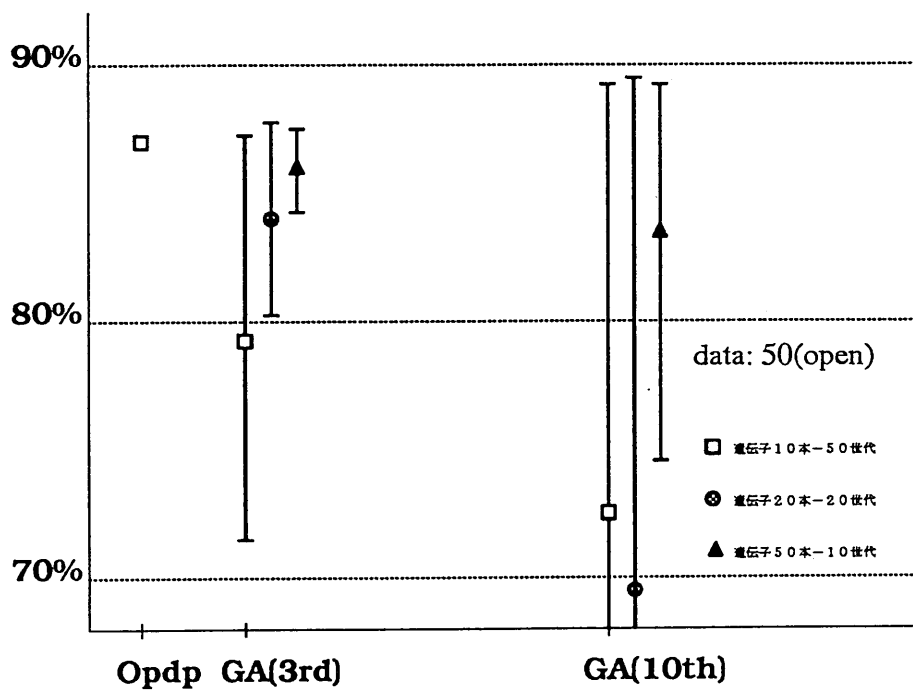


図 11: 遺伝子数、世代と認識率との関係 (抜粋)

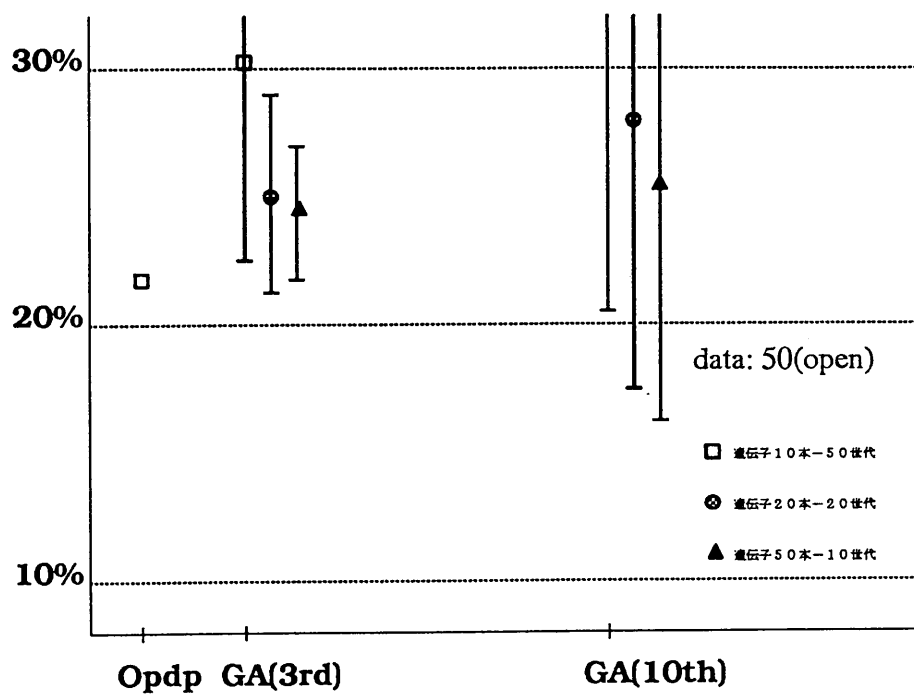


図 12: 遺伝子数、世代と誤検出率との関係 (抜粋)

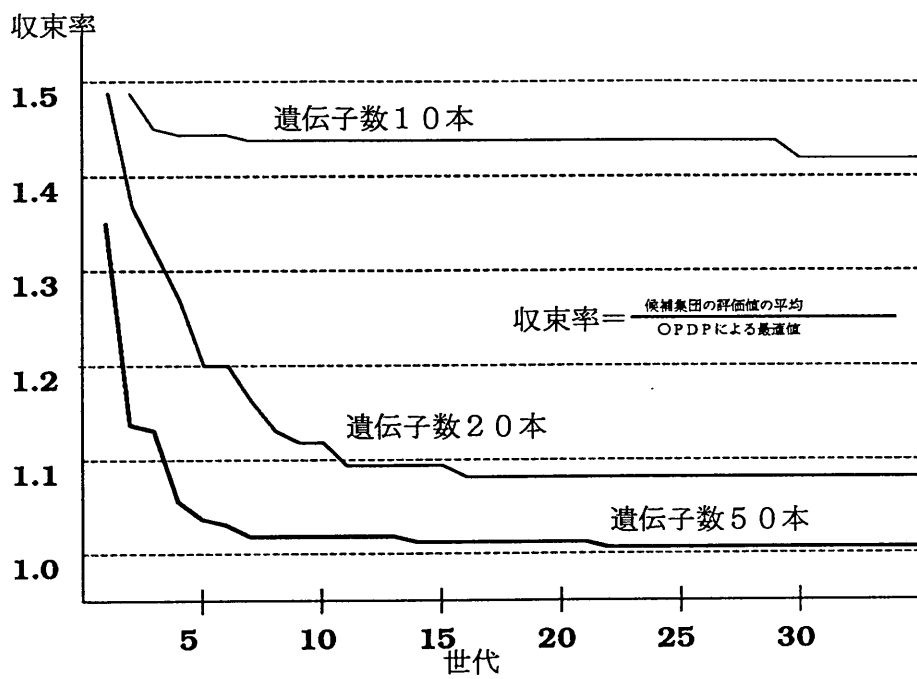


図 13: 遺伝子数と収束性との関係

句境界検出率

	第3位候補	第10位候補	One-Stage DP 法
Open data	87.17%	89.27%	86.50%
	85.58%	83.13%	
	83.96%	73.89%	
Closed data	86.89%	89.14%	85.68%
	85.02%	82.80%	
	82.95%	73.95%	

誤検出率

	第3位候補	第10位候補	One-Stage DP 法
Open data	23.34%	18.14%	25.77%
	26.07%	27.04%	
	28.65%	36.28%	
Closed data	26.27%	20.35%	29.28%
	28.91%	29.46%	
	31.44%	38.88%	

最良の候補のみの句境界検出率/誤検出率 候補全体の平均の句境界検出率/誤検出率 最悪の候補のみの句境界検出率/誤検出率

表 2: Open data と Closed data の句境界検出実験

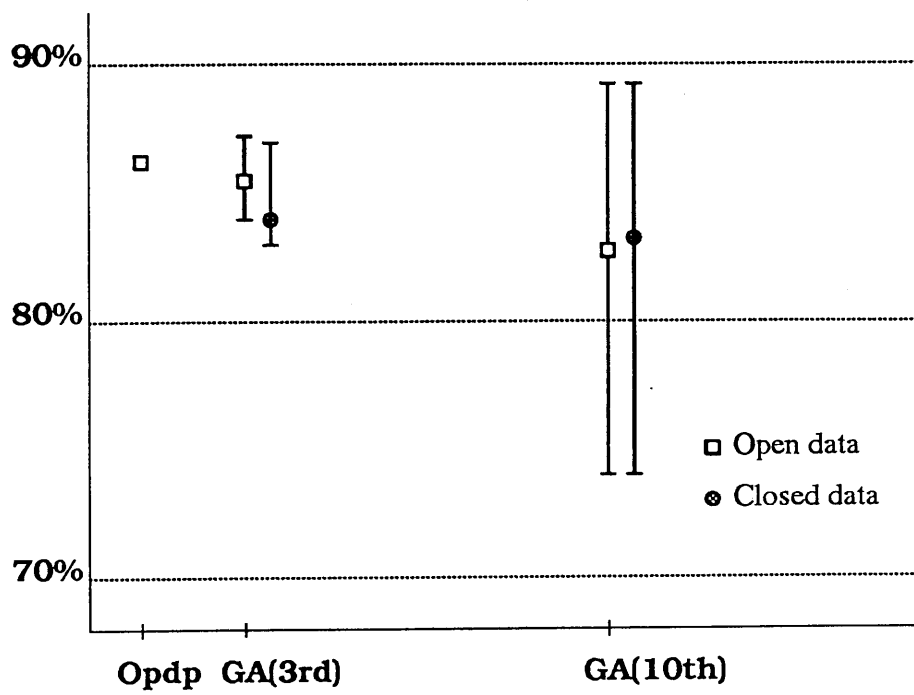


図 14: Open data と Closed data の句境界検出率

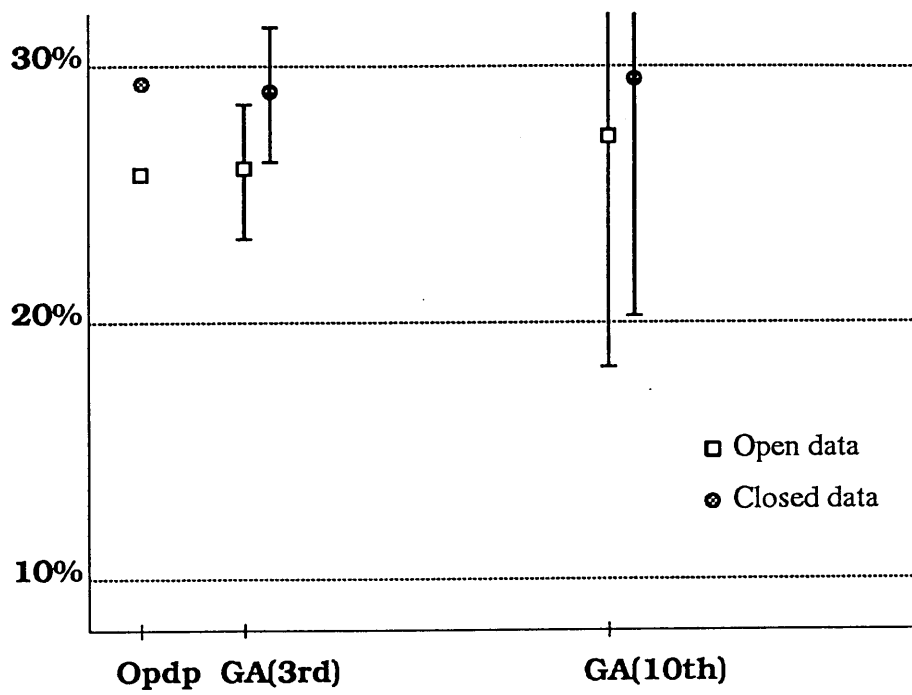


図 15: Open data と Closed data の誤検出率

4.4.1 実験方法

50 文章を遺伝子数 50 本 - 10 世代の GA 法で、100 位候補までとり、3/10/30/100 位候補までの句境界検出率を調べる。

4.4.2 結果・考察

結果を表 3、図 16、17 に示す。候補を増やすと、句境界検出率は向上し、100 位候補までとった場合、句境界検出率 91.98% という高い数字が得られた。しかしその反面、候補数が多くなるほど、句境界候補全体の評価も悪くなり、ほとんどの句境界候補が One-Stage DP 法の結果を下回る。つまり、候補を増やしても、良い評価の候補より、悪い評価の候補が多く出てくるのである。今回は句境界の正解が視察により与えられていて、その正解と候補とを比べることにより検出率を算出しているのので、多くの候補の中より最も良い候補を選び出せるが、一般には候補の情報のみで最終的な句境界候補を選ばねばならないので、候補数を増やすことが、即、句境界検出率の向上にはつながらないと思われる。

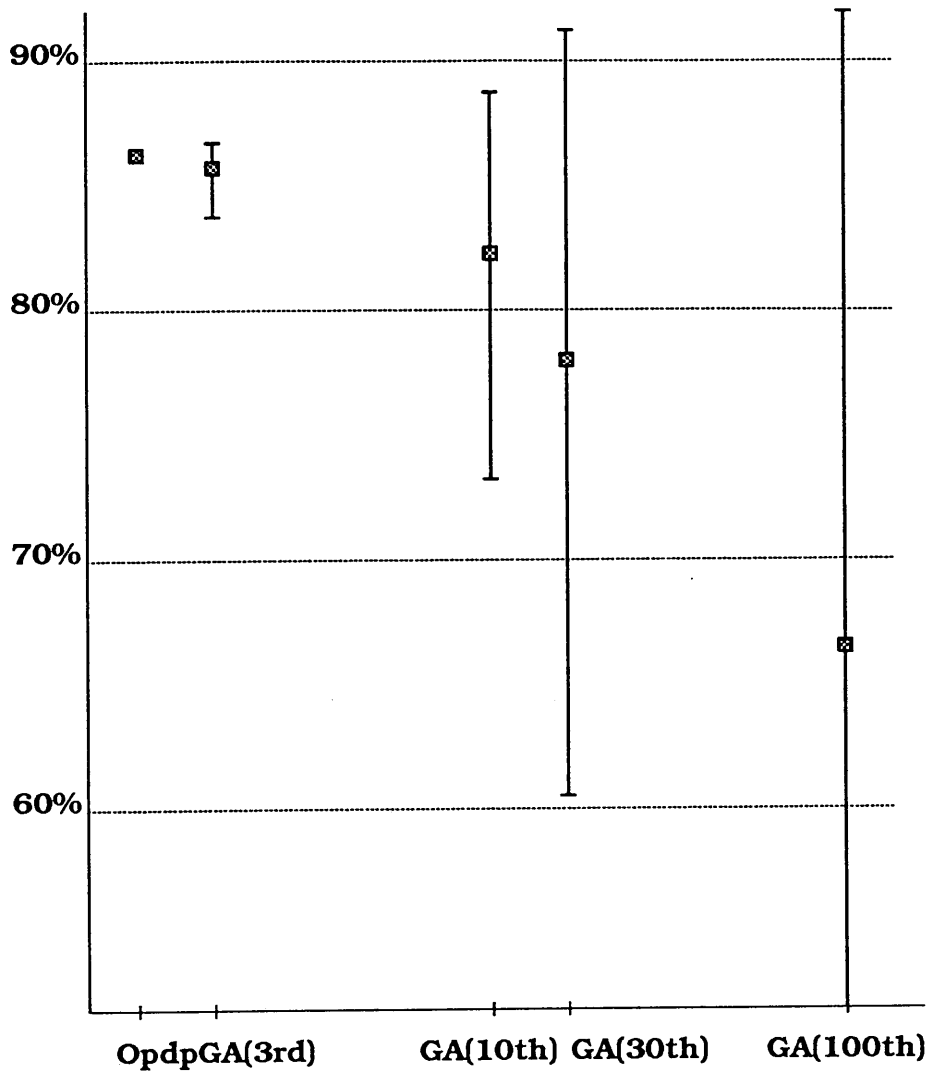


図 16: 100 位候補までの句境界検出率

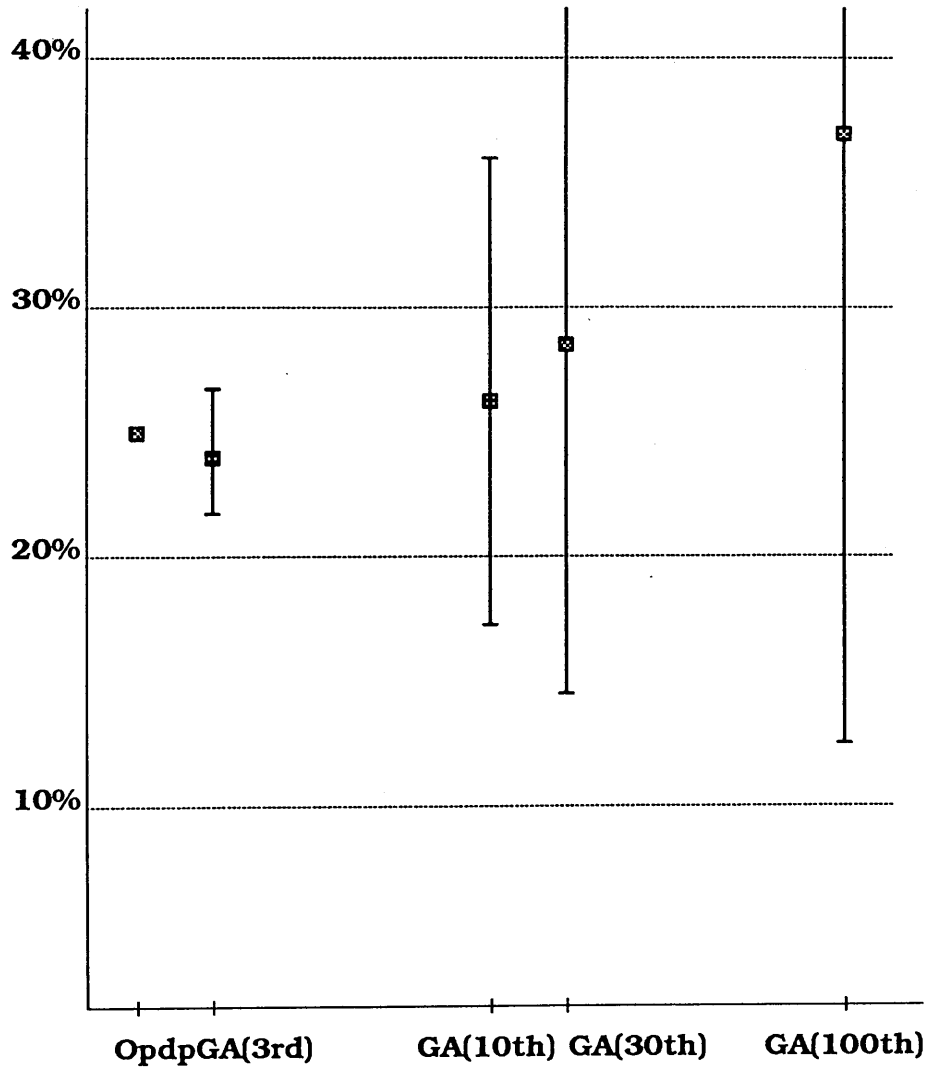


図 17: 100 位候補までの誤検出率

句境界検出率

	3 位候補	10 位候補	30 位候補	100 位候補	One-Stage DP 法
句境界検出率	86.79%	88.91%	91.27%	91.98%	86.32%
	85.61%	82.38%	77.86%	66.58%	
	83.96%	73.34%	60.14%	31.60%	
誤認識率	21.69%	17.21%	14.38%	12.50%	25.00%
	24.06%	26.24%	28.52%	36.94%	
	26.65%	36.04%	46.93%	67.92%	

最良の候補のみの句境界検出率/誤検出率
 候補全体の平均の句境界検出率/誤検出率
 最悪の候補のみの句境界検出率/誤検出率

表 3: 100 位候補までの句境界検出率

5 結論

5.1 結論

GA 法という最適化のアルゴリズムを利用することによって、以前に提唱されている One-Stage DP 法を上回る句境界検出率をあげることができた。しかし、たしかに、句境界候補集合の中には、One-Stage DP 法よりよい句境界検出率を持つ句境界候補がある反面、One-Stage DP 法よりも悪い句境界検出率を持つ句境界候補がそれ以上に多くあらわれてしまう。これは、本研究では触れていない、複数の候補からどうやって最終的な句境界候補を選出するかという問題にも、深く関わってくると思われる。

5.2 今後の課題

- 複数候補からの最終候補選択の方法

本研究では触れていないが、複数の句境界候補集団からどうやって最も句境界検出率の良い候補を選ぶか？

- 評価関数 DP 距離の改良

評価関数を DP 距離としたときの最適解が、One-Stage DP 法であるので、DP 距離を評価関数としたアルゴリズムが One-Stage DP 法より大幅にすぐれているとは考えにくい。DP 距離とは独立な特徴量を用いて作成した句境界候補集団と照らし合わせて最終候補を決定出来ないか？

謝辞

本研究を行うにあたり、多大な御指導、御助言、を戴いた阿曾弘具教授、堀口進助教授に心より感謝いたします。また、下平博助手には、本研究全般にわたり、親身な御指導を賜りました、ここに深く感謝いたします。さらに、有益な御示唆、御助言をいただいた中井満氏に感謝の意を表します。最後に、日頃より数多くの御指導、御討論、御協力を賜りました丸岡(東)研の皆様に感謝いたします。

参考文献

- Hermann Ney : “The Use of a One-Stage Dynamic Programming Algorithm for Connected Word Recognition”, IEEE ASSP-32, 2, pp.263-271(1984-04)
- R.Schwartz and Y-L.Chow : “The N-Best Algorithm An Efficient and Extact Procedure for Finding the N Most Likely Sebtebce Hypotheses”, Proc. ICASSP90, pp.81-84(1990)
- 下平博、木村正行、嵯峨山茂樹:” ピッチパターン連続整合による連続音声のセグメンテーション”、信学技報 SP90-72、pp.33-40(1990.12)
- 和田健之助:” 遺伝的アルゴリズムと機械の進化”、数理科学、1990.10
- 宮沢丈夫:” 遺伝的アルゴリズムと最適化問題”、ASCII, Vol 15, #6、1991.6