

グラフ探索アルゴリズムに基づく手話認識方式

竹内 勝、櫻井彰人、大淵康成

RWCP¹⁾新機能日立-i²⁾研究室

意味に関する事前知識を用いることにより、手話認識システムの認識精度を向上させることが可能である。本稿では、意味情報を利用した手話認識方式に関して、理論的基盤を与えることを目的とし、手指動作に関する形状情報の処理過程を経路探索問題の解法アルゴリズムとして、また、発話内容に関する意味情報の処理過程を被覆木探索問題の解法アルゴリズムとして表現する手法を考案している。さらに、この表現の下で、形状情報と意味情報を統合する手法を検討している。

Sign Language Recognition using Graph-Search Algorithms

Masaru Takeuchi, Akito Sakurai and Yasunari Obuchi

Novel Functions Hitachi Laboratory 1, RWCP

Japanese Sign Language (JSL) has few case markers, few words that represent voice, tense and aspect, and relatively free word order in a sentence; that is, it is syntactically for less restrictive than oral language. Thus, integration of semantic information is necessary for machine recognition of JSL. In this paper, an integration method based on graph-search problems, which were the shortest path problem and the minimum spanning tree problem, was presented.

1. はじめに

手話は手指動作のみならず、手指動作の様相や表情・口形などを用いるマルチモダル性をもつ情報伝達手段である。日本語では「は」や「を」などの助詞を用いて、文の中での単語の文法上の役割を表現する。一方、手話では「眉上げ」、「うなずき」といった動作で、文法に関する情報を表現する[1]。このため、計算機が、手指動作のみで手話を認識する場合、文法に関する情報を用いることが出来ず、これを補うために意味に関する知識を用いることが必要となる。既に原理実験では意味に関する知識を利用した手話認識方式は有望であるとの感触を得ている[2]。本稿は、手指動作の形状情報と発話単語の意味情報を利用した手話認識方式に関してを理論的基盤を与えることを目的とし、次節以降で、手話認識問題とその処理過程を、グラフ理論における探索問題とその解法アルゴリズムとして表現する方法を述べる。

2. 意味情報を用いた手話認識方式と原理実験 手話認識システム[3]にセンサーグループを用いて、 例えば

1) RWCP: Real World Computing Partnership
(新情報処理開発機構)

2) 株式会社 日立製作所 中央研究所内

「私は咳の酷い風邪を引いた。」

という発話内容を表す手指動作

<咳する, 多い, 咳する, なる, 私>

を入力すると、手話単語認識装置は連続的に手話単語を認識し、図1に示す手話単語束とよぶデータを出力する。単語束において、横方向は時間経過を示す。縦方向は標準パターンとの距離を示し、上方ほど手話認識装置が参照する標準パターンとの距離が小さいことを表す。単語束上のラベル付きの線分を切片とよぶ。切片はラベルで示す単語が線分左端の時刻から右端の時刻まで、縦方向の位置で示す標準パターンとの距離で、手話単語が認識されていることを示す。

手話単語系列生成問題は単語束から、実際に発話されたと考えられる手話単語系列を決定する問題である。形状情報のみから手話単語系列を決定する場合、比較的標準パターンとの距離が小さく、発話区間全体を覆い、互いに重複の少ないという条件の下で切片を選択するものとする、図1の例では、

<口, なる, 私>

という、発話内容とは異なる単語系列が決定される。これは「口」という手指動作と「咳」という手指動作が類似しているため起こる誤認識である。

意味情報は意味共起高階グラフを用いて表現する。ここで、グラフは二項関係の図式表現である。例えば図1の「多い」と「咳する」の間の辺は2つ

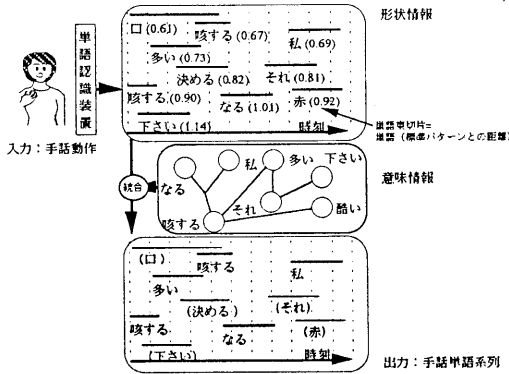


図1:意味情報を用いた手話認識方式の概念図

の単語が意味的にまとまりが良いことを示す。図では省略したが、辺には重みとよばれる数値が付加されており、意味的なまとまりの良さの程度を表している。

高階グラフはグラフを一般化した多項関係の図式表現である。図1では「咳する」、「なる」、「私」の3単語の意味的なまとまりが良いことが表されている。「多い」、「咳する」や「咳する」、「なる」、「私」といった手話単語の組み合わせが、意味的にまとまりが良いという意味情報を用いることによって、

<咳する, 多い, 咳する, なる, 私>

という正しい手話単語系列を決定することが可能となる。

原理実験の概要を示す。原理実験は57単語からなる32例文について行った。意味共起高階グラフは、コーパスから統計的に抽出せず、単語の二項組、三項組、各々について意味共起の重みを主観的に与えることにより構築した。

重みの与え方は、以下に示す(1)~(3)の過程を多数回、繰り返すことによって決定する。

- (1)複数の単語の組を計算機スクリーン上に提示する。
- (2)提示した単語の組の順位を主観的に決定し、並べ替える。
- (3)上位に位置する単語の組の重みを増加させ、下位に位置する単語の組の重みを減少させる。

手話単語系列の決定は、以下に示す(1)~(4)の手順で手話単語系列の解候補の生成と評価を繰り返し、生成した複数の解候補を比較し、最適な評価値をもつ解候補を手話単語系列として出力する。

- (1)手動作に関する形状情報を用いて手話単語系列の解候補を生成する。
- (2)生成した解候補について、形状情報に関する評価値を与える。
- (3)生成した解候補について、意味共起高階グラフを用いて意味情報に関する評価値を与える。
- (4)適当な加重定数の下で(2), (3)の評価値の線形加算結果を解候補の評価値とする。

32例文を6回ずつ発話した192例に関して、上記、手話単語系列の決定方法で手話文認識を行い、一位認識率を比較した。その結果、上記、(3)の評価値、つまり、意味情報を使用した場合は、使用しない場合に比べて、誤認識を51%削減した。

意味情報を用いる手話認識方式では、手指動作に関する形状情報は単語束により表現されている。また、発話内容に関する意味情報は重み付きのグラフにより表現されている。原理実験では、形状情報に

関する評価値と、意味情報に関する評価値を別個に求め、それらの加重加算を行うという方式で2種類の情報を統合している。

次節以降では、両者の情報処理過程のデータ表現を共通化し、意味情報を用いた手話認識方式を、グラフ探索問題を基盤として、定式化する。また、手話認識システムで取り扱う単語数が増加すると、認識に要する処理時間が増加する。本稿では言及しないが、本稿の定式化は、既存のグラフ探索問題の高速な解法アルゴリズムの利用や新たなアルゴリズムの考案による処理時間増加へ対処を目的としている。

第3節では、両者の情報処理過程のデータ表現を共通化するために、統計的な解析により、単語束で表現されている形状情報を意味情報と同じ形式のグラフで表現する方法を与える。さらに、グラフで表現した形状情報から、意味情報と統合可能な情報を導出するための解析を行う。第4節では、意味情報を表すグラフの構築手法を与える。また、意味情報の処理過程を構築したグラフの上での被覆木探索問題の解法アルゴリズムとして、2種類の形式で定式化する。第5節では、第3節の解析をもとに意味情報と統合可能な手指動作の形状に関する情報を定義し、第4節で扱う意味情報と、統合する方式を与える。この統合により、形状情報と意味情報が併せて単一のグラフで表現される。このグラフに、第4節のアルゴリズムを適用し、手話単語系列を決定する。

3. 手動作情報の処理のグラフ表現

本節では、両者の情報処理過程のデータ表現を共通化するために、統計的な解析により、単語束で表現されている形状情報をグラフで表現する方法を与える。さらに、グラフで表現した形状情報から、意味情報と統合可能な情報を導出するための解析を行う。

なお、以下、本稿では高階グラフではなく、二項関係を表す通常のグラフのみを扱うものとする。また、グラフは全て有向グラフであるものとする。

3.1. 形状モデルと単語束グラフ

単語束をグラフで表現するために、ラベルと重みの付いたグラフ

$$G = \langle V, E, L_V, W_E, f_V, f_E \rangle \quad (1)$$

を定義する。Gを単語束グラフとよぶ。Vは集合であり、グラフ理論ではその要素を頂点とよぶ。単語束グラフでは、頂点の集合は3種類の頂点により構成される。

$$V = V \cup \{v_h, v_t\} \quad (2)$$

とする。Vの各々の頂点は単語束の切片と一対一に対応する。また、 v_h, v_t は各々、文頭と文末を表す頂点であり、手話単語系列の始点と終点を明示するためのものである。 L_V は頂点に与えたラベルの集合である。頂点のラベルは対応する切片の単語ラベル、または、「文頭」、「文末」を与える。

$$f_V: V \rightarrow L_V \quad (3)$$

は、頂点とラベルの対応を与える写像である。Eは集合であり、グラフ理論ではその要素を辺とよび、切片間の二項関係の有無を表す。辺には重みwが与えられ、その値域は W_E である。

$$f_E: E \rightarrow W_E \quad (4)$$

は辺と重みの対応を与える写像である。

単語束グラフは頂点の集合とそのラベル、辺の集合とその重みを構成することにより決定される。

手話認識システムに手話単語が入力されると手話単語認識が連続的に行われ、単語束が出力される。このとき、単語束の各切片に対応する、単語ラベルのついた複数の頂点と文頭、文末の頂点を与えることにより、単語束グラフの頂点の集合とそのラベルは容易に構成できる。

単語束グラフを決定するためには、その辺集合と、(4)式の辺の重みを定めればよい。この重みが有限であるとき、その辺は辺集合の要素と見なす。辺の重みは、単語束の切片について、その単語束上の位置関係に関する統計的解析結果から決定する。

単語束の統計的性質の解析方法を述べる前に、単語束の切片とその位置関係について検討する。

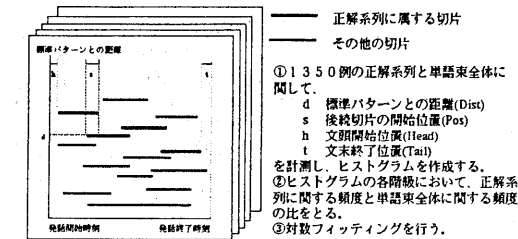
単語束の1つの切片を選択すると、その標準パターンとの距離 d (図2)を知ることができる。切片はその標準パターンとの距離 d が小さいほど正解系列に含まれている可能性が高い。ここで、発話内容を表す単語系列を正解系列とよぶ。切片の標準パターンとの距離 d に応じて、その切片が正解系列に含まれている可能性の強さを、多数の単語束に関しての統計的な解析により、標準パターンとの距離 d の関数として求めることが出来る。

一方、同時に複数の単語が発話されることはないため、正解系列に含まれる連続する2つの切片において、先行する切片の発話終了時刻と後続する切片の発話終了時刻の差 s (図2)については、発話時間に関して重複が少ないものと考えられる。さらに、1つの手話文の発話はスムーズに行われ、言い淀みや間といった、無発話の時間区間、つまり、長い空白がないことを仮定すれば、先行する切片の発話終了時刻と後続する切片の発話終了時刻の差 s は0に近い値をとるものと考えられる。標準パターンとの距離の場合と同様に、先行する切片の発話終了時刻と後続する切片の発話終了時刻の差 s についても後続切片が正解系列に含まれている可能性の強さを、先行する切片の発話終了時刻と後続する切片の発話終了時刻の差 s の関数として求めることが出来る。

なお、本稿では、このような切片のパラメータの値に対して、その切片が正解系列に含まれる確率を与える関数を確率モデルとよぶ。

以上の検討に基づき、統計的性質の解析方法と確率モデルとよぶ関数の構成方法を述べる。

切片 i とその標準パターンとの距離 $d(i)$ 、切片 i に対するその後続切片 j の開始位置 $s(i, j)$ 、さらに、切片 i に関する文頭の開始位置 $h(i)$ と切片 j に関する文末の終了位置 $t(j)$ を、複数の例文に対して計測し、単語束全体に含まれる全ての切片についてのヒストグラム



例文135文×発話10回=単語束1350例

図2: 形状情報に関する統計解析方法の説明図

と正解系列に含まれる切片のみについてのヒストグラムを作成する。そして、作成したヒストグラムの各階級において、単語束全体に関する頻度に対して、正解系列に関する頻度の比を計算する。標準パターンとの距離の場合、この値は、任意に選択した切片の階級値に応じて、その切片が正解系列に含まれる確率を与える。一方、後続切片の開始位置の場合、この値は、任意に選択した2つの切片についての階級値に応じて、先行する切片が正解系列に含まれているという条件の下で、後続する切片が正解系列に含まれる確率を与える。

この頻度の比に対数当てはめを行い、標準パターンとの距離 $d(i)$ 、後続切片の開始位置 $s(i, j)$ 、文頭の開始位置 $h(i)$ と文末の終了位置 $t(j)$ について、パラメータの値に対して、その切片が正解系列に含まれる確率を与える、確率モデルとよぶ関数を作成する。これらを各々、距離モデル P_{Dist} 、位置モデル P_{Pos} 、文頭モデル P_{Head} 、文末モデル P_{Tail} とよぶ。また、総称して形状モデルとよぶ。

図3に135例文を各々10回発話した1350文について、作成した各モデルを図示する。ここで、図3の「dim」で示す、対数当てはめを行う際の各モデルの次数は情報量基準(AIC)を用いて決定している。

単語束グラフの頂点と頂点 j を結ぶ辺の重み $w(i, j)$ は、図3の距離モデル P_{Dist} と位置モデル P_{Pos} の独立性を仮定し、

$$f_E = w(i, j) = -\log(c \times P_{Pos}(s(i, j)) \times P_{Dist}(d(j))) \quad (5)$$

で与えるものとする。ここで、 c は頂点の出力辺全体に関して、(5)式の \log の内部の和が1となるための正規化のための定数、

$$c = 1 / \sum_j P_{Pos}(s(i, j)) \times P_{Dist}(d(j)) \quad (6)$$

である。この正規化により、(5)式の \log の内部は、頂点 i が選択されたという条件の下で、頂点 j が選択される条件付き確率を表している。

また、(5)式において、 $-\log$ を作用させるのは、一般的にグラフ理論では重みの和の最適化を扱うため、確率の積演算を重みの和演算に変換するためのものである。

なお、

$$P_{Pos}(s(i, j)) \times P_{Dist}(d(j)) = 0 \quad (7)$$

の場合は、(5)式が定義できないため、

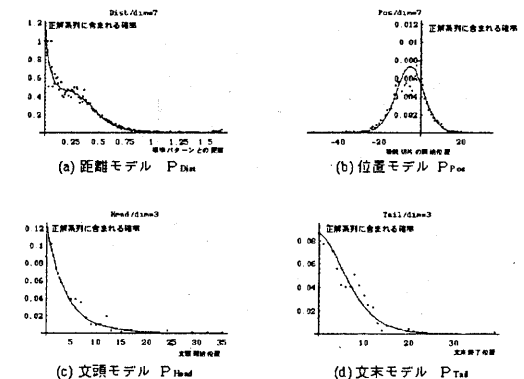


図3: 形状モデル-切片が正解系列に含まれる確率

$$w(i,j) = \infty \quad (8)$$

とし、この場合、グラフ理論では一般に、辺は存在しないものと見なす。

文頭、文末についても、同様の処理を行い、図4に示す単語束グラフを得る。ここで、 $w(i,j)$ を第 i,j 成分とする行列を W で表し、単語束グラフの隣接行列とよぶ。

3.2. 単語束グラフと解候補の生成

手話単語系列は単語束グラフ上では、図4の太線で示すように、文頭を表す頂点を始点、文末を表す頂点を終点とする経路として表現される。さらに、形状モデルの下で、文頭から文末へ至る全ての経路の空間において、経路、

$$\text{Path} = \langle i_0, i_1, i_2, \dots, i_n \rangle \quad (9)$$

が正解系列である確率は、(5)式、(6)式より、

$$P_{\text{Path}} = \exp[-\{w(i_0, i_1) + w(i_1, i_2) + \dots + w(i_{n-1}, i_n)\}] \quad (10)$$

で与えられる。このとき、その経路長を w_{Path} とすると、

$$P_{\text{Path}} = \exp(-w_{\text{Path}}) \quad (11)$$

となる。従って、単語束の形状情報のみから、正解である確率が最も高い手話単語系列を求める問題は、単語束グラフ上で文頭を表す頂点を始点、文末を表す頂点を終点とする最短経路問題として表現される。また、原理実験の場合のように単語束情報をもとに手話単語系列の解候補を複数列挙するには、第 k 最短経路生成アルゴリズム(4)とよばれる、経路長の短い順に経路を列挙する経路探索アルゴリズムを単語束グラフに適用すれば、正解である確率が高い順に手話単語系列を列挙することが出来る。

3.3. 形状情報のマルコフモデル

ここでは、単語束グラフとマルコフ連鎖の関係を述べ、意味情報と統合する手指動作に関する形状情報を導出するための検討を行う。

単語束グラフの辺の重み $w(i,j)$ に対して、

$$p(i,j) = \exp(-w(i,j)) \quad (12)$$

但し、

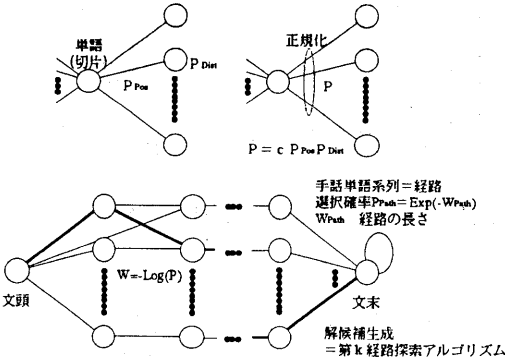


図4: 単語束グラフ説明図

$$w(i,j) = \infty \quad (13)$$

の場合は、

$$p(i,j) = 0 \quad (14)$$

とすると、(5)式、(6)式より

$$\sum_j p(i,j) = 1 \quad (15)$$

が成立する。さらに、単語束グラフの文末を表す頂点に重さ1のループ追加し、重みを $w(i,j)$ から $p(i,j)$ に書き替えることで、図4の単語束グラフはマルコフ連鎖の状態推移図としての諸性質を満たす。ここで、 $p(i,j)$ を第 i,j 成分とする行列を P で表し、単語束グラフの状態推移行列とよぶ。

状態推移行列の各成分 $p(i,j)$ は切片 i が正解系列のある位置に含まれていた場合、切片 j が正解系列のその次の位置に含まれる確率を示す。

状態推移行列 P の冪乗行列 P^n の i,j 成分、 $p^n(i,j)$ は、切片 i から n 回の状態推移で切片 j へ至る確率を表す。つまり、 $p^n(i,j)$ は、切片 i が正解系列のある位置に含まれていた場合、切片 j が正解系列において、 n 単語後に含まれる確率を表す。なお、

$$P^n(i,j) = 0 \quad (16)$$

の場合は切片 i から切片 j へは長さ n の経路で到達可能でないことを表わしている。

第5節では $p^n(i,j)$ を用いて意味情報と統合するための手指動作に関する形状情報を定義する。

4. 意味情報の処理のモデル

本節では、事前知識としての意味情報を表すグラフの構築手法を与え、意味情報の処理過程を、構築したグラフの上での被覆木探索問題の解法アルゴリズムとして、2つの形式で定式化する。

4.1. 意味共起グラフの構築

原理実験では主観確率を用いて意味共起グラフを構築した。これに対して本稿では、次項で述べる述語を中心とした意味情報の処理のモデルの下でコーパスから意味共起グラフを構築する手法をのべる。

意味共起グラフ構築のため、日本語コーパスにおいて、

- (1) 述語の頻度
- (2) 主語-述語の関係、主語-目的語の関係、修飾語-被修飾語の関係の頻度
- (3) 係り受け関係における係り先の単語に関して、係り元がない場合の頻度

を計数し、手話辞典で日本語単語を手話単語に変換し、その結果をもとに手話単語に関する頻度情報を計算し、記録する。

図5の例では、

「私は熱が高く、鼻水が酷い、風邪を引いている」

という例文から自立語を抽出し、その係り受け関係を調べると、図5の手話辞典の左側に挙げる係り受け関係が得られる。ここで、修飾語と非修飾語の関係は1例のみを示し、残りは省略している。手話辞典を用いて、これらの係り受け関係を手話単語に変換したものを、図5の手話辞典の右側に示す。このような手話単語の組に関する頻度を、多数の例文について計数する。

意味共起グラフとしては、停止頂点を使用しない場合と使用する場合の2種類のグラフを係り受け関係に関する頻度情報をもとに構築する。停止頂点と

例文「私は熱が高く、鼻水が多い、風邪を引いている。」

自立語「私」「熱」「高い」「鼻水」「多い」「風邪」「引く」

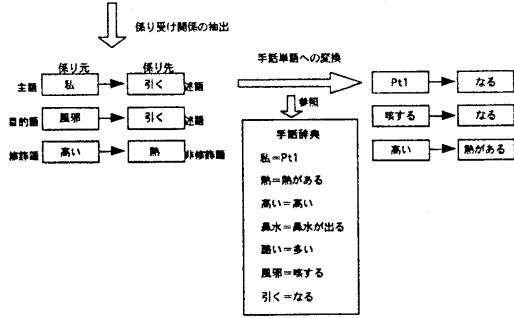


図5:係り受け関係の抽出方法の説明図

は(3)の係り元がない場合を表現するための頂点である。

第5節では3つの形状情報と意味情報の統合手法を述べる。このうち、停止頂点をを用いない統合手法においては、単語束から手話単語系列を決定する過程で、第3節の第k経路探索アルゴリズムに基づく、形状情報による解候補の生成過程が必要となる。現在、第k経路探索アルゴリズムは自作のものを用いて、研究を進めているが、その計算量は単語数に対しての4乗のオーダーであり、処理時間の上での負荷が高い。第k経路探索アルゴリズムの使用を避けるため、停止頂点を導入した。

停止頂点を使用しない意味共起グラフの頂点の集合は根頂点とよぶ1つの頂点と手話単語のラベルが付いた複数の頂点により構成される。頻度情報のうち(1)の述語の頻度を根頂点から各手話単語への辺に付加する。また、(2)の各単語間の係り受け関係の頻度が0以上の場合は、係り先の手話単語から係り元の手話単語への辺に付加する。そして各頂点の出力辺全体に関する重みの総和で各々の辺の頻度を正規化し、その値 q を

$$m = -\log(q) \quad (17)$$

で変換し、辺の重みとする。

停止頂点を使用する意味共起グラフの頂点の集合は根頂点とよぶ1つの頂点と手話単語のラベルが付いた複数の頂点、および、同数の停止頂点により構成される。頻度情報のうち(1)の述語の頻度を根頂点から各手話単語への辺に付加する。また、(2)の各単語間の係り受け関係の頻度が0以上の場合は、係り先の手話単語から係り元の手話単語への辺に付加する。さらに、(3)の頻度は手話単語に対応する頂点から各々の停止頂点への辺に付加する。そして、停止頂点以外の各頂点についてその出力辺全体に関する重みの総和で各々の辺の頻度を正規化その値 q を(16)式で変換し、辺の重みとする。

4.2. 文生成モデル

上記2つの意味共起グラフに対して、各々、意味情報の処理モデルを定義する。意味情報の処理モデルを文生成モデルとよぶ。両者ともグラフ探索問題における最小被覆木探索問題の解法アルゴリズムであるPrimのアルゴリズム(4)を基に構成する。

停止頂点を使用しない文生成モデルは、手話単語の集合が与えられた場合、その手話単語に含まれる頂点と根頂点が張る停止頂点を使用しない意味共起

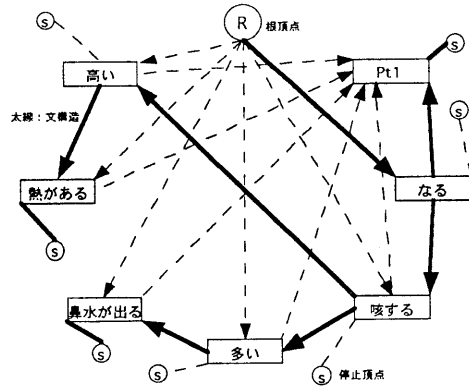


図6:意味情報の処理の結果生成される文構造の例

グラフの部分グラフに対して適用される。根頂点を出発点として、最初、集合 V にはこの頂点だけを入れ、残りの(部分グラフの)頂点すべてを集合 U に入れる。生成過程の各ステップでは、集合 V の頂点から集合 U の頂点への辺の中から重みが最小のものを選び、この辺を文を構成する係り受け関係の一部として確定するとともに、この辺の集合 U 側の頂点を集合 V へ移動する。これを全部の頂点が集合 V へ移動するまで繰り返す。

停止頂点を使用する文生成モデルは、停止頂点を使用する意味共起グラフの任意の部分グラフに対して適用される。停止頂点を使用しない文生成モデルと同様に根頂点を出発点として、最初、集合 V_0 にはこの頂点だけを入れ、残りの(意味共起グラフ全体の)頂点すべてを集合 U に入れる。また、集合 V_1 は空集合とする。生成過程の各ステップでは、集合 V_0 の頂点から集合 U の頂点への辺の中から重みが最小のものを選ぶとともに、この辺の集合 U 側の頂点が停止頂点である場合は、この辺の両側の頂点を集合 V_1 へ移動する。停止頂点でない場合は、この辺を文を構成する係り受け関係の一部として確定し、集合 U 側の頂点を集合 V へ追加する。この過程を集合 V_0 が空集合になるまで繰り返す。

図6は、停止頂点を使用する意味共起グラフの例である。同図において、停止頂点を使用する文生成モデルで得られる文構造を太線で示す。

以上2つの文生成モデルでは、生成した文の意味情報に関する評価値を、生成した文を表す木の辺集合 E_B について、その辺 e_k に付与された重み m_k の総和を E_B の辺の数で割った値、

$$g(E_B) = \frac{\sum_{e_k \in E_B} m_k}{\#(E_B)} \quad (18)$$

で与えるものとする。ここで、 $\#$ は集合の濃度を表すものとする。

5. 形状情報と意味情報の統合

ここでは経路探索問題の解法アルゴリズムとして表現された形状情報の処理過程と、被覆木探索問題の解法アルゴリズムとして表現された意味情報の処理過程の、3種類の統合方法について述べる。

第1の統合手法は、原理実験の場合と同様に、形状情報から手話単語系列の解候補を生成し、解候補の形状情報に関する評価と意味情報に関する評価を個別に行い、これらの評価値を用いて手話単語系列を決定する手法である。

アルゴリズムの概要を示す。手話単語束から単語束グラフを構成する。次に、単語束グラフに対して、第k経路選択アルゴリズムを用いて、K個の手話単語系列の解候補を生成する。ここで、Kはあらかじめ定められた解候補の生成個数である。生成した各々の解候補の形状情報に関する評価値を経路長WPathで与える。次に、その解候補を構成する手話単語が張る意味共起グラフの部分グラフを抽出する。そして、この部分グラフについて、停止頂点を使用しない文生成モデルを用いて文構造を決定し、意味情報に関する評価値を(18)式の $g(E_B)$ で与える。そして、解候補の評価値を形状情報と意味情報に関する評価値の加重加算、

$$c_1 W_{Path} + c_2 g(E_B) \quad (19)$$

が最小となる解候補を手話単語系列として決定する。

第2の統合手法は第1の統合方法と同様に形状情報から第k経路選択アルゴリズムを用いて、解候補の生成を行う。ただし、第1の統合方法において、解候補を構成する手話単語が張る意味共起グラフの部分グラフを抽出するとき、意味共起の強さを表す部分グラフの辺の重みに、後述する形状情報から導出される評価値を加重加算する。そして、停止頂点を使用しない文生成モデルを用いて文生成を行い、解候補の評価値を $g(E_B)$ で与え、この値が最小となる解候補を手話単語系列として決定する。

この統合手法で、加重加算する形状情報の評価値の計算方法の概要を示す。3節の単語束の状態移行行列 P^n の冪乗行列 $P^{n,i}$ は、 i, j 成分、 $p^n(i, j)$ は、切片 i から n 回の状態推移で切片 j へ至る確率を表している。この統合手法と第3の統合手法では、それらの有限の加重和、

$$P^*(i, j) = \left(\sum_{n=0,1,\dots,N} \alpha_n P^n \right) (i, j) \quad (20)$$

により、切片の形状情報によるままとりの良さを定義し、この値を意味情報の重みに加重加算する。(20)式は、切片 i が正解系列に含まれている場合、切片 j がその正解系列の切片 j より後ろの部分に含まれる程度を表す。ここで、単語束の状態移行行列 P は文末ノードが唯一の吸収状態であり、この状態の生起確率が1となる最小のべき乗回数を N とする。

また、(20)式は、意味的な係り受け関係がある2つの手話単語(切片)が手話文(手話単語系列)において、連続して発話されない場合、それらの形状情報に関する評価値を間接的に与えているためのものである。例えば、

＜熟、高い、鼻水、多い、咳する、なる、私＞
(熱が高く鼻水の酷い風邪を引いた)

の場合、係り受け関係にある「高い」と「咳する」の間の形状情報に関する評価値を間接的に与えている。

第3の統合手法では、手話単語束全体に現れる手話単語が張る意味共起グラフの部分グラフを抽出する。このとき、第k経路選択アルゴリズムによる解候補の生成は行わない。次に、第2の統合手法と同様に、意味共起の強さを表す部分グラフの辺の重みに、(20)式の形状情報に関する重みを加重加算する。そして、この部分グラフについて、停止頂点を使用する文生成モデルを用いて、文構造を決定する。最後に、生成した文を構成する手話単語を手話単語束に基づき整理し、手話単語系列として決定する。

この統合方式は、複数の解候補を生成し、その評

価値を比較し、手話単語系列を決定するのではない。手動作の情報と、意味情報を意味共起グラフの上に統合し、そのグラフ全体に関する最小被覆木問題を、ただ一度、解くことで手話単語系列を決定している。

6. 結言

本稿では、意味情報を用いる手話認識方式の処理過程を、グラフ理論における探索問題の解法アルゴリズムとして表現する方法を与え、この表現の下で、形状情報と意味情報を統合する手法を検討した。ここで、グラフ理論における探索問題とは、例えば、最短経路探索問題はグラフ上の2点を結ぶ最小の長さの経路を見つける問題で、最小被覆木探索問題はグラフの全頂点を結ぶ大きさが最小の木を見つける問題である。

主要な結果は次の3件である。

- (1)意味情報を用いる手話認識方式のうち、手指動作に関する形状情報の処理過程を経路探索問題の解法アルゴリズムとして表現する手法を考案した。また、マルコフ連鎖として表現する手法も考案した。
- (2)意味情報を用いる手話認識方式のうち、発話内容に関する意味情報の処理過程を被覆木探索問題の解法アルゴリズムとして表現する手法を考案した。
- (3)形状情報の処理過程のマルコフ連鎖表現から導出される情報を意味情報に統合し、被覆木探索問題の解法アルゴリズムを用いて手話単語系列生成を行う手法を考案した。

本稿では、二項関係を表す意味共起グラフを用いて、意味情報の処理過程のモデルを構築した。今後の課題としては、原理実験で用いた、多項関係を表す意味共起高階グラフを包含する意味情報の処理過程のモデルの構築を行うことである。また、近似アルゴリズムや確率アルゴリズムを用いたアルゴリズムの高速化の検討も行いたい。さらに、早急に評価用データを作成し、アルゴリズムの性能評価実験を行うことも必要である。

また、性能評価実験で良好な結果が得られた場合は、実データへの適用を行う。このためのコーパスとして、既に5000文程度の病院での会話例文を用意している。ただし、4節で述べた述語を中心とした係り受け関係を抽出するためには、信頼性の高い構文解析ツールを作成することが必要となる。

参考文献

- (1) 米川明彦: 手話言語の記述論的研究, 明治書院(1984)
- (2) 竹内勝, 櫻井彰人: 意味的共起関係を用いた手話認識方式, 自然言語処理106-9(1995)
- (3) 佐川浩彦, 他: 圧縮連続DP照合を用いた手話認識方式, 信学論J77-D-II, 4(1994)
- (4) 石畑清: アルゴリズムとデータ構造, 岩波書店(1989)