

連続音声認識と物語理解を目指した言語処理について

重永 実

中京大学情報科学部
470-03 豊田市貝津町床立101

後続の単語を予測し、物語の事象間の連鎖関係を推論できるように、種々の知識と予測機能を汎用的な形で連続音声認識システムと物語理解システムに組み込んできた。音声認識システムでは汎用的な意味情報と拡張された格構造の四つ組のセットを使って、名詞と述語の予測が行われる。さらにシステムは単語の連想機能をも持っている。物語理解システムでは種々の文脈情報を使って、文脈解析が行われ、事象間の連鎖関係を求め、エピソード・ネットワークを形成し、物語のあらすじを生成することができる。

Linguistic Processors for Continuous Speech Recognition and Story Understanding

Minoru Shigenaga

School of Computer and Cognitive Sciences,
Chukyo University
Tokodate 101, Kaizu-cho, Toyota, 470-03 Japan

In order to predict following words and infer relation among events in a story, the author has been trying to implement various knowledge and predictability, in universally applicable forms, into both speech recognition and story understanding systems. In the speech recognition system, using generalized semantic information and a set of four terms of extended case structures of predicates, prediction of nouns and predicates is performed. It has also associative function. In the story understanding system contextual analysis is performed using various contextual information, and the system can infer relation among events, compose an episode network, and give outlines of stories.

1. はじめに

計算機に連続音声認識させようとする、調音結合の影響や発音が曖昧になり易いことなどによる識別音素の不完全さや単語或いは句(文節)境界が不明なことが、文章の認識を甚だしく困難にしている。しかし、我々は通常全ての音素を必ずしも正しく識別しているとは考えられず、持っている知識による補完により、全ての音素を正しく聞き分け無くとも認識、理解しているものと思われる。一方、文章の理解にも色々なレベルがあり、対象や目的によっては、キーワードや表層的な意味だけで十分目的を達成できる場合もある。しかしこの場合でも、我々が必要な知識を持っているから可能になっていると思われる。そこでこの困難を軽減するため、音声認識の領域では、タスクを限定し、確率モデルによる HMM (隠れマルコフモデル) や trigram などの統計的処理の使用により対処している¹⁾。また自然言語処理では、意味・文脈解析には C/D 理論、スクリプト、プラン、ゴール²⁾ などの提案による予測機能の導入が計られているが、正しい文字列が入力されるため、一文の中で後続の単語の予測を考慮したものはないようである。というより、その必要はなく、寧ろ主語などの省略への対応や照応関係の解明を始めとして、曖昧性の解消、限定された話題に関する談話理解などに努力が払われている。

また、従来、連続音声認識をも含めて、自然言語の構文解析には通常文脈自由文法やそれを発展させたものが使われているが、文の複雑さが増してくると、文法の記述や管理がかなり面倒で、追加修正も容易ではない。また、意味情報や制御機構などを途中に入れることも容易ではなく、まず形態素解析、構文解析、意味解析、文脈解析と進められるのが普通である。しかし我々は必ずしもこの順序に処理しているとは考えられない。即ち、我々が文章を読んだり、会話をする場合、一文の中でも、次に続く文に対しても、常識を含めた多くの知識、経験に基づくスクリプト的な記憶や連想機能などのような文脈に関連した知識や機能を駆使して、予測し、検証して、理解していくものと考えられる。そしてその理解の程度は我々が持っている知識の量と質によるものと考えられる。

筆者らは人間の情報処理に少しでも近いと思われる音声認識や物語理解を目指して、音声認識システム(語彙 1018 単語)では、後続単語の予測能力を高め、一文中では形態素解析、構文解析、意味解析を一体として処理していくことを試みてきた。ここではまず、構文情報を時枝文法の司と辞の考えを取り入れた構文状態遷移ネットワークにより表現し、拡張した格文法を汎用的な形で使用している。構文情報のネットワーク表現によるオートマトン制御は極めて単純明快であり、途中で制御機構を入れたり、意味情報を挿入することも容易であり、複雑な複文で無ければ句の間の係り受け関係も容易に決定し

うる。かつ、人間にとって理解し易く、拡張、変更も容易である。意味処理に関しては、各品詞間の修飾関係を汎用性を持たせた意味属性の形で表して、品詞間の接続関係の制約や名詞句や述語の予測機能の充実に努めている。現在のシステムも基本的には以前の報告^{1,2,3)}と大差ないが、より変化に富んだ文章構造への対応や語彙の増加などに対処するため、構文情報を拡張し、意味情報の充実に努め、格構造の四つ組を準備するなど、言語およびタスクに無関係で、汎用的な予測能力を高めるための種々の改善を加えてきている^{6,7,8)}。かつ、名詞と動詞の原形のみを入力により文章を構成できる⁹⁾。

物語理解では文脈解析に重点をおき、C/D 理論、スクリプト、プラン、ゴールなどのほか、新たに“概念ユニット”を導入するなどして、物語中の事象間のより深い連鎖関係を求めて、より充実した物語表現(エピソード・ネットワーク)を実現し、このネットワークを使用し、物語のあらすじの生成を試みている。

2. 連続音声認識システムにおける言語処理

2.1 構文情報^{4,8)}

2.1.1 句の定義と修飾機能

時枝文法に従って、単語をその構文上の役割に着目し、体言や用言のような素材や概念を表すもの(概念語)と、助詞や助動詞のように、概念語に付属して、それに特定の構文的な機能を与えるもの(機能語)の2種類に大別する。両者は結合して(結合したものを句と定義する)文中に現れ、後続の語句を修飾する。この構文上の修飾機能は

- (1) 連体修飾機能 (f1): 後続の体言を修飾する機能。
- (2) 連用修飾機能 (f2): 後続の用言を修飾する機能。
- (3) 完結形式 (f3)

の3種類に大別される。つまり文は(1)、(2)の2種類の修飾機能を持つ句が、次々と語句間の係り受け関係を決定してゆき、最後に完結形式で終わり、ある意志表示が完結するものと言える。図1に句構造やその修飾関係の1例を示す。図中に点線で示すように、句の定義により、単位の句は集まって、より大きな句 P8, P9, P10 を形成していく。これらの修飾関係を構文木で表すことも容易にできる。

2.1.2 構文状態遷移ネットワーク

日本語の語句間の隣接関係を構文状態遷移ネットワーク(図2)で表現している。図中の大きな円で囲まれた SOC (節の始端)、SOP (句の始端)、SVP (動詞の始端)等は、句を識別した時に到達する状態を示し、小さな円で囲まれた N (名詞)、V (動詞)、A (形容詞)等は概念語を識別した時に到達する状態を示す。長円で囲まれた SOS、EOS はそれぞれ文の初めと終わりを示している。また大きな円(または SOS)から小さな円へ向かう矢印につけられた記号は概念語の品詞の略記号(表1)であ

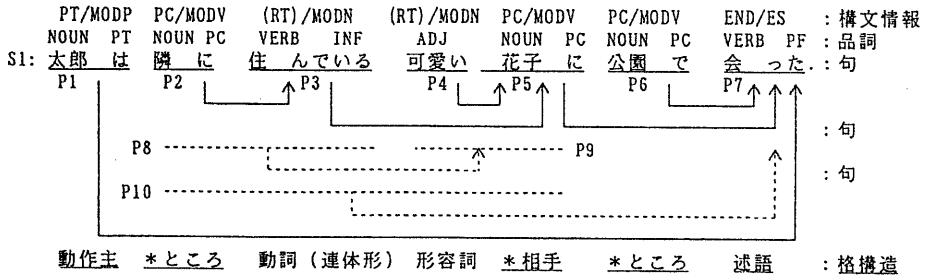


図1 単語、句構造とそれらの間の修飾関係

表1 単語の分類

品詞	略記号	単語例	備考
名詞	NOUN	山, 絵, 彼	代名詞を含む
形容詞	ADJ	大きい, 綺麗だ	形容動詞を含む
動詞	VERB	行く, 食べる	
連体詞	PREN	その, 大きな	
副詞	CADV SADV DADV	もし, たとえ はっきり, 早く とても, 大変	条件 動作, 状態 程度
接続詞	CONJ	すると, そして	
助詞	PP PT PC PF	の, と は, も が, に ので, か	連体修飾関係 副助詞 格助詞 接続助詞, 終助詞
助動詞	AUX	です, ない	
活用語尾	INF	(行) きました (綺麗な) ので (学生) です	述語(動詞, 叙述 形容詞, 叙述名詞) の語尾を含む

表2 活用形の分類

活用形略記号	説明
RY	連用修飾形
END	終止形
RT	連体修飾形
COND	条件形
IMP	命令形
INT	疑問形
PAUSE	文の中断を示す形

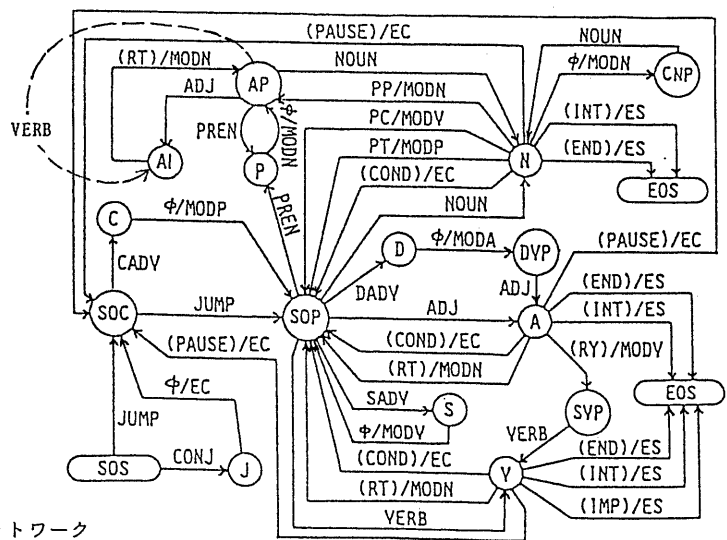


図2 構文状態遷移ネットワーク

る。小さな円から大きな円（または EOS）に向かう矢印につけられた記号は機能語に対する構文情報を示し、X/Yの形をしている。そして①その機能語が活用しない場合、Xはその機能語の品詞の略記号（表1参照、ただし空記号φを含む）である。またYは構文関係を示し、連体修飾関係を示すMODN（名詞の修飾）、連用修飾関係を示すMODV（動詞の修飾）、MODA（形容詞の修飾）、MODP（述語の修飾）がある。またEC、ESはそれぞれ節の終わり、文の終わりを示す。②機能語が活用する場合にはXはその活用形を示す。この場合のXは表1のAUXまたはPF、あるいはそれらが結合したもので、括弧でくくられており、その略記号の意味は表2に示されている。また、Yは①の場合と同じである。なお、助動詞と助詞PFsは図2には陽に現れず、1382個の活用語尾と共に括弧の中のXに埋め込まれている。JUMPは無条件ジャンプを示す。

構文解析は通常文脈自由文法などに拠って、非終端記号を介して探索が行われる。しかしここでは、あるノード（状態）にあるとき、まず構文的、意味的予測が行われて、予測された単語群とのみ照合が行われる。従って、形態素解析、構文解析、意味解析と順次行うのではなく、構文解析、意味解析、予測、連想、形態素解析が同時に行われる、というより単語や句は最後に決定されるという特徴を持っている。かつ、語順が自由な言語に対して有用であり、文脈自由文法などに比べて表現が極めて単純であり、視覚的にも常にその全容を把握し得、処理できない構文に遭遇しても、他の遷移に影響を与えたり、矛盾を生ずることなしに、新しいアークやノードを追加して対処することが容易にできる利点がある。

例えば文章入力システム¹³⁾（語彙5982単語）では、図2のノード“AP”に破線で示すアークを付加することにより、5才児程度の子供を対象に書かれた世界昔話集のうちの約50物語に含まれる説明文（会話文を除く）2675文章も処理できるようになっている。

さらに、文末からの解析や文中の任意の単語（句）からの解析（島駆動）にも容易に使用できる。

2.2 単語の意味表現^{5,6,8)}

(1) 名詞

各名詞に通常複数個の階層構造を形成した意味属性（*をつけている）を与えている。

(2) 形容詞の属性

形容詞を意味的に25種類に分類して、その各々の意味を形容詞の属性と呼ぶことにする。そして名詞の意味属性にはそれと結合しうる形容詞の属性を記述しておく。例えば属性“気持ち”をもつ形容詞には“楽しい、優しい”などがある。名詞毎にこれらの属性を入れておくのではなく、その意味属性にこれらの属性を与えておく。例えば*人という意味属性には属性として“気持ち、性格”等を入れておく。すると“優しい彼”は許されるが、

“優しい果物”は許されず、“形容詞+名詞”の意味的チェックと名詞の限定に使用可能である。

(3) 述語（動詞、叙述形容詞、叙述名詞）の格構造

(1)で述べた名詞の意味属性を利用して、動詞、叙述形容詞、叙述名詞の意味を拡張した格構造を使って表現している。格になりうる句は一般には格助詞を含んでいるが、例えば<*人-は>、<*人-も>のように格助詞以外の助詞を含む名詞句をも含めて、広い意味で格と呼んでいる。

2.3 述語及び名詞句の予測^{6,7,9)}

2.3.1 2句間の関係による名詞などの限定

単語のもつ意味属性を使って、汎用的な形で、次の場合について後続の名詞などを限定している。

「形容詞+名詞」、「連体詞+名詞」、「'その'+名詞」、「名詞+'の'+名詞」、「名詞+'と'+名詞」、「名詞+'や'+名詞」、「[連体詞+名詞]」、「[動詞の連体形+名詞]」、「副詞+動詞」、「[連体詞間の接続関係]」、「副詞や助詞の呼応」等。

2.3.2 格構造による述語（動詞、叙述形容詞、叙述名詞）と名詞句の予測—前向き予測と後向き予測

名詞句と述語とは非常に密接な関係を持っており、先行の名詞句を認識できれば、後続の述語を予測できる（前向き予測）。逆に述語を先に認識できれば、後続の名詞句を予測できる（後向き予測）。筆者らのシステムでは句境界情報の利用により、語尾の述語を抽出し、認識することが可能であり、構文状態遷移ネットワークを文末から逆向きに遷移することも容易にできる。

なお、連体形の動詞の格構造を利用して後続の名詞を予測しうる。

2.3.3 名詞句の二段階予測

前向き処理の場合でも名詞句を予測することができる。すなはち、一つまたは二つの名詞句が認識されれば、先ず述語を予測する（前向き予測）。するとたとえその述語を認識しなくても、予測された述語の格構造を使って、先行の名詞句を予測することができる（後ろ向き予測）。

2.3.4 前向き予測のみによる名詞句と述語の予測—格構造の四つ組

上記の予測の一元化、高速化を考える。そのため先ず、述語の格辞書を使って、“名詞の意味属性 + 格助詞 + 格の種類 + その格を持つ述語”の四つ組を集めた予測用知識ベースを構築しておく。するとこの知識ベースを使って(1)名詞に続く格助詞の予測、(2)名詞句からその後に続く名詞句の予測、(3)名詞句からの述語の予測を行うことができる。

なお、文末からの処理では最初に述語を認識して、この知識を使っていることになる。

2. 4 連想機能^{14,9)}

上記の種々の情報のほかに、我々は連想機能を持っていて、既に認識した文章や単語から文脈やある概念、単語などを連想して、文章の理解を助けていると考えられる。そこで約 180 名の大学生を対象に連想単語の調査を行い、連想単語辞書を作成し、候補単語の絞り込みに利用している。

2. 5 結果⁹⁾

2.5.1 予測の効果

表 3 に「おじいさんは山へ柴刈りに行きました」が認識されてゆく途中、それまでの認識結果が正しい場合、各々の時点で予測された後続の名詞、述語の数を示す。特に述語の予測が合理的に行われていることがわかる。

2.5.2 名詞、動詞からの文章生成

表 4 (a)に名詞と 1 個の動詞の原形のみを与えた時に生成される単文の数を示す。助詞の使用制限が難しく、意味の無い文が生成されている。(b)は助詞を格助詞と「のは」に限定した場合である。ただし、実際には色々な語尾変化が許されるので、この数倍の数が現れる。また、この実験はシステムの意味情報のチェックにも役立つ。

3. 物語理解を目指して^{10,11,12)}

物語理解の研究で特に先駆的なものは R.G.Schank らの²⁾の研究であり、理解における記憶の仕組みや階層構造に関連した様々なレベルの知識の必要性を明らかにした。また、一般に物語の理解の程度を示すには、(1)物語の内容についての質問に答える、(2)物語のあらすじを生成できる、(3)物語の言い替えができる、(4)類似した物語を想起できる ことなどの側面が挙げられるであろう。

筆者らは人間が物語を理解する過程を、“読み手が既に持っている実世界に関する常識や知識、物語の次の展開についての予測などを用いて、物語内の事象を意味的に結び付ける過程である。”と見なし、(1)物語中に陽に述べられていない事柄を様々な種類、様々なレベルの推論規則を用いて物語表現に組み込む、(2)物語中の事象間の関係を求める、(3) (1), (2)より得られる物語の展開をエピソードのネットワークとして記憶の中に残すことに主眼を置いて、物語理解システム JStory の構築に努めてきた。つまり、物語理解を事象のネットワークを作成する過程であると見なし、種々の物語理解の側面をエピソード・ネットワークの利用により実現しようとしている。そこで JStory では、スクリプト、プラン、ゴールや TAU¹⁵⁾などをはじめ、新たに提案する“概念ユニット”¹⁰⁾という概念の連鎖関係を汎用的に表す知識などを使った推論機能を利用して行間を埋めたりすることにより、より充実したエピソード・ネットワークを作り出そうと努めている。このエピソード・ネットワークを利用することにより、物語中の事象間のつながりの流暢さ、物

語全体の構造の明確な表現、質問に対する応答の解り易さ、大事な事象の抽出の適切さ、他の物語との類似点の抽出のし易さなどの特徴を持つと共に、物語のあらすじを生成できる。

3. 1 JStory の概要

本システムは構文解析部、意味解析部、文脈解析部、質問文理解部、文生成部、あらすじ生成部から成る。構文解析部（音声認識システムのそれとほぼ同じ）では、ローマ字で分かち書きされた入力文から句構造を抽出し、構文木を作る。意味解析部では、構文木を基にして、係り受けなどの処理を行い、入力文の意味表現をつくる。文の意味表現は格構造、法情報を取り出したものと、CD (Conceptual Dependency Theory) 表現からなる。文脈解析部では、意味解析部で出来た意味表現から適切に様々な文脈知識に起動をかけ、事象間のつながり（エピソード・ネットワーク）を作る。質問文理解部では、質問文の類別や何を答えれば良いのかを判断し、質問文に対する答えを見つけるために、エピソード・ネットワークまたは辞書知識の中を探索する。文生成部では、質問文に対する応答文やあらすじ文を生成する。あらすじ生成部では、キー・センテンスの抽出と、あらすじの図式を使い、あらすじ文を生成する。

3. 2 構文解析部

入力文がローマ字で分かち書きされており、辞書内容が幾分異なるほかは、音声認識システムのそれとほぼ同じである。

3. 3 意味解析

構文解析で得られた句表現を構成する全ての句について句の性質を解析し、構文木に表された係り受け関係を調べて、意味的整合性のない構文木を排除し、文の意味表現をつくる。意味表現は基本的には格構造、CD 表現、アスペクトなどの法情報を使用している。動詞のプリミティブには Schank らのものの他に、「何かを作り出す、または表現する」ことを示す product や「ある出来事が起こるまで待つ」ことを示す waitなどを付加している。述語辞書には CD 表現のプリミティブ名と概念格、その概念格に対応する格構造の格の名前が記述してある。格構造の要素を利用し、CD 表現を完成する。もしも埋め込み文がある場合には、その埋め込み文が修飾する名詞の属性に埋め込み文の意味表現を記入し、かつ、それをイベントとして陽に出している。CD 表現を構成し終わると構文情報、格構造、CD 構造、述語辞書項目番号などを一つのフレームにする。このフレームをイベントと呼ぶ。

3. 4 文脈解析

文脈解析部では、意味解析部で事象を意味表現（イベント）に変換した文と前文までのすべての意味表現（イ

イベント)との間に何らかの関係を求め、イベントとイベントとをエピソード・リンク(表1参照)でつなぎ、物語(文章)全体としての構造を抽出する。物語中のイベントとイベントとを意味的に接続してできるネットワークは、物語の進行の軌跡を残していることになる。このネットワークをエピソード・ネットワーク(以後EPNと略記)と呼ぶ。文脈解析での目標は、大筋として、このEPNを作り上げることにある。そのため、名詞や指示詞の同定などの照応関係の処理、複数名詞の処理、省略された格の推定などの処理をして、同定した名詞に必要な属性を追加すると共に、(1)因果規則、(2)スクリプト、(3)プラン、(4)ゴール、(5)概念ユニット、(6)TAU、(7)物語の時空マップ、(8)物語構造知識¹⁶⁾、(9)その他の推論規則のような様々な知識を利用して、(1)物語中のイベントとイベントの間の関係を求める、(2)物語文章の中で陽に述べられていないイベントを推論し、EPNに組み込む。などのことを行っている。

3.4.1 概念ユニット (Conceptual Unit)

スクリプトは、記憶ないしは知識の表現に対して有用な構造である。しかし物語の流れにはスクリプト的概念には似ているが、スクリプトの処理能力を越えた、明かに一つの巨大な因果連鎖により結びついている状況が存在していることが多い。例えば、「恩返し」とか「脅迫して何かをさせる」といったものであり、これらは関係する動作主間での価値観や、そのときの状況などにより様々な手法で行なわれ、スクリプトのような標準的シーンというものが存在しない。しかし、動作主間での関係の概念が連鎖関係を保持しているのが特徴である。この知識には以下のような利点がある。(1)物語の中心概念に着目することが出来る。(2)物語内の文と文の結合性の強弱まで把握することが出来る。(3)要約時に使用することが出来る。(4)長期記憶として残るのはこの部分である。(5)物語間の類似性を捉えることが出来る。この知識を概念ユニットと呼ぶことにする。これは他の文脈解析の知識の処理とも結びついていて、柔軟性のある知識となっている。概念ユニットとしては動作主間でポジティブ、ネガティブな行為が行なわれると、それにより以後どの様な行為が行なわれるのか、それによりどの様な感情の想起が行なわれるのかに着目したもの(pos-effect, neg-effect)。出発-行動-帰路といった大きな一連の動作を表現したもの(nanika-sini-iku)。脅迫することにより相手に何かをさせるといったことを表現したもの(threaten)などが現在準備されている。例えば、広義の「恩返し」を表現した概念ユニット pos-effect は「人1が人2に好印象を与えたら、人2は人1にとってポジティブ(positive)なことをする」という概念関係を表現している。

3.4.2 TAU¹⁵⁾

M. Dyer は失敗を中心にした一連の行動に関連した人間の知的作業をTAU (Thematic Abstraction Units) という知識構造を用いて計算機上で実現している。本研究においては、否定文と fail リンクが連鎖しているイベントの説明、理由、およびそのイベントに関連する処理に関して、TAU の考え方を使っている。例えば、「笠地蔵」では、おじいさんは最初6人のお地蔵様に笠をかぶせたいと思うが、笠が一枚足りず、代わりに手ぬぐいをかぶせてあげたという部分である。この「ところが、笠が一つ足りません」という入力文に対して使用している。その結果、「笠を数(6-5)個欲しい」というゴールをおじいさんは想起する。そしてこの文以後のゴール・プランの処理により「おじいさんは笠の代用品となる手ぬぐいを(6-5)個準備する」となる。

3.4.3 その他の推論規則

・相手に対する感情についての推論

C1 : 「X は Y をかわいそうに(かわいそうのたぐいの感情)思った」

C2 として、Y が何かゴールの達成に失敗、または Y が精神的、肉体的打撃を受けた事象を見つけ、C1 と C2 を initiate リンクでつなぐ。

・登場人物の性格の解析

登場人物の形容詞修飾からリンクを張ることも行っていて、例えば「欲張りな」に対しては“何かを手にいれたい”というゴールを通常の人よりも重要に思っている人の性格である。従って

・X が欲張りな性格を持つなら、X のゴールとして「何かを手にいれたい」というゴールを推論する。

・X が欲張りな性格を持ち、X 以外の人物 Y が何らかの利益を得るなら、X が Y と同じ行動を取るとき次のことを推論する。

C1 : X は Y が利益を得たことを知っている。

C2 : X が Y と同じ行動を取るのには、X が Y の得た利益を欲するためである

そこで、C1 と C2 を motive リンクでつなぐ。

などの推論規則が用意されている。

3.4.4 物語の時空マップと物語構造

物語の時空マップとは、物語中で起こった出来事の間所関係や時間関係を表現したもので、EPN 中の各イベントに対するポイントを含んでいる。マップ情報は例えば“雨が降る”事象が起これば、雨を避けるゴールは建物の中にいない登場人物に対してのみ適用するというように、推論規則を適用する人物を限定したり、また時間や場所に関する質問の処理に利用している。

一方、JStory では物語を “物語 = 設定部 + 展開部”¹⁶⁾ という構造で見ている。

3. 5 物語の意味表現—エピソード・ネットワーク

物語をどのように表現すれば良いのかという問題は、人が物語を理解した後、物語を記憶の中でどのように表現しているのかという問題が解明されていない現段階では、明確な答を得ることができない問題である。本研究では、構文・意味・文脈解析の結果、物語内で生じた事象に相当するイベントと、事象間を関連付けている連鎖に相当するリンクとによって成り立つネットワークを物語の内部表現としている。そのネットワークを EPN、EPN 内のリンクをエピソードリンク、EPN 内のノードのうち入力文を表現したノードを入力イベント、システムが推論して新たに作成して行間を埋めているノードをデータイベントと呼んでいる（文献 10,12 参照）。

3. 6 物語間の類似性について

人間が物語を読み進めている時に、今読んでいる物語と似た筋の展開を持った過去に読んだ物語を思い出すことがある。そしてその思い出した類似の物語の筋の展開から物語の次の展開を予想したりする。もしも複数のテキスト間で類似した筋の展開があるならば、それらのテキストの物語表現の中にも色々な形で反映があって然るべきであろう。そこで JStory で処理した物語名とその物語で使用された概念ユニット名を調べると、表 4 のようになり、「こぶとりじいさん」「舌きり雀」「笠地蔵」が動善の点で、特に前二者はさらに懲悪の点でもよく似ており、人や動物が協力を (coop 概念ユニット) 大きなかぶを抜く「大きなかぶ」は他の 3 者と全く異なる展開であることがわかる。

3. 7 質問応答¹⁰⁾

表層的な質問にだけでなく、かつ深い理由も加味して答えられるが、どのような質問にでもとはいかない。

3. 8 あらすじ生成^{11)・12)}

あらすじ (要約) 生成の手法はこれまでにいくつか提案されており、目的や対象により使い分ける必要がある。特にその対象を物語とする場合には、まず物語を深く理解し、文の因果関係を十分考慮して、あらすじを構成する文章間の連鎖関係を維持することが必要であると考えられる。しかしこれまでにこのような意図をもって具体的にあらすじを生成したものはないと思われる。

そこで本稿では JStory が物語を理解した結果として生成する事象のネットワークに基づいて、ボトムアップ的にキーセンテンスを抽出する手法を採用した。それは前述のように、事象のネットワークは物語を理解し、文章内の各事象間の因果関係を表現したものであるから、このネットワークに基づいてキーセンテンスを抽出することにより、物語の展開 (筋) を追跡し、それを構成するキーセンテンス間の連鎖関係 (原因—結果、理由—行

動) を維持したあらすじの生成ができると考えられるからである。文献 17) でも同じ考えの下に種々試みているが、本稿ではその結果を勘案、補強し、かつ文献 17) では十分な対応ができなかった異なるタイプの物語 (例えば “大きなかぶ”) にも対処できるように改善されており、物語の筋を維持し、文と文との間の連鎖関係を保持したあらすじを生成することができるようになっている。そして不十分な実験ではあるが、人間の指摘とかなり同じようなキーセンテンスを抽出できていることを確かめている。

さて、一般に物語は登場人物などの行動の記述が始まるまでの設定部とそれ以後の展開部とに分けられるから、

あらすじ = 物語設定部 + キーセンテンスの連鎖と定義する。そして、出入りするリンクの本数が多いノードは、それだけ物語の内容と強く結びついて、重要なノードであると判断される。何本以上のリンクを持つノードを重要なノードとするかは次式で決定する。

(四捨五入 (一つのノードに連鎖するリンクの本数の平均)) + 1

この本数以上のリンクを持つノードをキーセンテンス抽出の基となるノードとする。従って本稿におけるあらすじ生成は重要な連鎖関係を維持しながら、キーセンテンスに相当するノードをエピソード・ネットワーク内の全ノードの中から抽出する作業となり、その出力形式としては、通常、抽出したノードの文章の連鎖となる。

生成したあらすじについては文献 12) を参照されたい。

3.8.1 あらすじの評価¹²⁾

人のあらすじ生成と比較してシステムの評価をするため、大学生 6 人 (質問 (3) のみ 8 人) を被験者として、次の実験を行った。すなはち被験者にまず、配布した四つの物語を読んで貰い、(1) 自由なあらすじの生成、(2) 大切な文及び部分の選択、(3) 物語中で用いられている文を使用してのあらすじ生成を依頼した。その結果、(1) では意外にもほとんどの人が物語文章中の文をそのままなぎ合わせたものをあらすじとしており、(3) による文章とほとんど同じであった。かつ文章の長さもほとんどのものが原文の半分を越えていた (本システムではほぼ半分以下に抑えている)。元の文章が子ども向きに簡略化されていることによるものと考えられる。(2) の結果から選ばれた文 (節) に含まれている重要事項を因子と呼び、5 人以上の人が選んだ因子を必須因子と名付け、(3) において 6—8 人の人があらすじに取り入れた因子を重要因子と名付けた。そして、これらの必須及び重要因子を本システムが生成したあらすじが含み、極めて重要でない因子を含んでいないかどうかを調べたところ、どの物語に対しても、20—30 個の因子のうち 1、2 の因子を除いてこの条件を満たしていた。従って、裏付けとしては不十分ではあるが、本論分で取り上げた四つの昔話に対しては、ほぼ人間と同じように重要な文

を取り上げてあらずじとしていると言えるであろう。満たされない主なものは、表層のみによるつながりしかない文が重要因子を含んでいる場合である。例えば「こぶとりじいさん」では、「その時、たくさんの鬼が現れました」という文は表層上のみによるつながりしかないため、EPN 上には出てこない。本システムではこのような前後と密な関係の無い文中の小設定は無視されるからである。しかし、この文をあらずじの中に取り入れている人が多くいた。今の場合後続の文から容易に推論可能なので、この文がどうしても必要だとは考えられないが、一般的には、文中の小設定をどのような場合にどの程度考慮に入れるべきかは、今後検討する必要があるだろう。

4. おわりに

音声認識システムにおいては意味情報の充実による単語の予測能力の強化、物語理解システムでは文脈解析の充実によるイベント間の連鎖関係の明確化に努めてきた。今後更に種々の知識を増強し、処理できる物語の数を増すと共に、両システムの結合が必要である。また、より高度の連想機能の導入と共に、文脈による後続の単語や文章の予測機能を高める必要がある。音声認識システムにおいても初期においては意味ネットワークを形成したが、各出力文章が正しくなければ使用できない。少なくとも意味的に誤りのない文章を抽出する(複数個ありうる)か、誤り修正機能を付与する必要がある。

これらの研究は筆者が山梨大学計算機科学科在籍中やその後における関口芳廣助教授を始め多くの研究室の学生諸君の努力の賜であり、改めて深謝します。

参考文献

- 1) 例えば Lee K.F. et al.: An overview of SPHINX speech recognition system. IEEE Trans. on ASSP, 38, 1, pp.35-45(1990).
- 2) 例えば Schank R.C. et al.: Scripts, Plans, Goals and Understanding, Lawrence Erlbaum Associates (1977).
- 3) M. Shigenaga et al.: Speech recognition system for Japanese sentences. Proc. COLING-80, pp. 472-479 (1980).
- 4) 関口ほか: 日本語文章の音声認識システムにおける構文情報について. 信学論, J65-D.6, pp.782-789 (1982).
- 5) 関口ほか: 日本語文章の音声認識システムにおける意味情報について. 信学論, J66-D.6, pp.629-636 (1983.6)
- 6) 関口ほか: 連続音声認識のための格構造を利用した述語予測. 電学論, 108C.10, pp.818-825 (1988.10).
- 7) 重永ほか: 連続音声認識システムにおける述語及び名詞句予測の可能性について. 信学論, J72-D-II, pp.1307-1312 (1988.10).
- 8) Shigenaga M. et al.: A Large Vocabulary Continuous Speech Recognition System with High Predictability. IEICE Trans., E 74, 7, pp.1817-1825 (1991).
- 9) 神部ほか: 単語予測機能を備えた不特定話者連続音声認識システム. 平成4年春季音講論, pp77-78 (1992.3).
- 10) 野崎ほか: 物語理解におけるエピソードネットワークの構築. 情報処理論文誌, 30, 9, pp.1103-1110 (1989).
- 11) 中澤ほか: エピソードネットワークを用いた物語のあらずじ生成. 情報処理論文誌 32.10, pp.1215-1224 (1991).

- 12) 中澤ほか: 物語のエピソードネットワークの構築とあらずじ生成, 中京大学情報科学部 TECHNICAL REPORT, No.91-2-04 (1992).
- 13) 石橋ほか: 語句間の関係を考慮した大語意連続音声入力システム. 信学技報, SP91-92 (1991.12).
- 14) 山口ほか: 大語彙連続音声認識システムにおける連想単語辞書の利用. 信学技報, SP90-76 (1990.12).
- 15) Dyer M.: "In-Depth Understanding", MIT press (1983).
- 16) Rumelhart D.E.: Notes on a schema for stories, in Bobrow D.G. and Collins A., Representation and Understanding Studies in Cognitive Science (1975).
- 17) 野崎ほか: エピソードネットワークを用いた物語理解. 情報処理学会, 知識工学と人工知能研究会資料, 51-4, pp.25-32 (1987).

表3 予測単語の数 (1)はスタート時点)

1) おじいさんは 2) 山へ 3) 柴刈りに 4) 行きました

句	名詞		述語		形容詞		動詞	
	無	有	無	有	無	有	無	有
1)	535	535	535	535	70	70	242	242
2)	535	256	79	40	52	30	217	116
3)	459	92	0	0	0	0	41	14
4)	245	52	0	0	0	0	8	2

(有/無: 連想情報の有無)

表4 概念語群のみによる生成文章の数

* 概念語群

- ① 三匹 小豚 家 作る
- ② チム 板切れ のごぎり 切る
- ③ おじいさん 山 しばかり 行く

(a) 助詞を限定しない場合

概念語群		出力文数		意味OK
		肯定文	否定文	
①	肯定文	6	5	9
	否定文	1	4	
②	肯定文	1	2	1
	否定文	1	2	
③	肯定文	3	3	3
	否定文	3	6	

(b) 助詞を格助詞と”の、は”に限定した場合

概念語群	①		②		③	
	肯	否	肯	否	肯	否
文の種類	3	3	3	3	16	16
全出力数	3	3	3	3	16	16
意味OK	3	3	3	3	16	16

表5 物語ごとの使用概念ユニット

物語名	使用概念ユニット
こぶとりじいさん	pos-effect, neg-effect, nanika-sini-iku, meet, threaten
舌切り雀	pos-effect, neg-effect, nanika-sini-iku, meet
笠地蔵	pos-effect, nanika-sini-iku
大きなかぶ	coop