

概念依存関係による概念素統合に基づいた音声理解

永井 明人、石川 泰

{nagai, yasushi}@media.isl.melco.co.jp

三菱電機株式会社 情報技術総合研究所

〒 247 神奈川県鎌倉市大船 5-1-1

発話表現の自由度が高い音声対話システムを実現するために、概念素に基づく意味理解方式の実用化検討を進めている。理解性能の向上にはタスク知識の積極的な利用が重要であり、このためには、タスク知識を一般的な言語知識と独立に扱う必要がある。そこで、本方式の言語的制約知識に関して一般的な言語知識とタスク知識との分離を試み、一般的な言語知識として導入した概念依存関係に基づいて、概念素間の部分依存構造を構築していく音声理解手法を提案する。さらに、語彙 1000 単語規模の鎌倉観光案内の自由な質問を対象とした音声理解実験を行ない、本手法の音声理解性能を評価した結果を報告する。

Speech Understanding Based on Integrating Concepts Related by Conceptual Dependency

NAGAI Akito and ISHIKAWA Yasushi

MITSUBISHI Electric Corporation

Information Technology R & D Center

5-1-1 Ofuna, Kamakura, Kanagawa 247, JAPAN

We have proposed concept-driven semantic interpretation method, which is in the improvement stage for practical application in order to realize such a spoken dialogue system that robustly understands various expression uttered by a naive user. For this improvement, to utilize domain knowledge is important, and system must have portability. This paper discusses generalization of the semantic interpretation method, and proposes the method which integrates concepts using general linguistic knowledge of conceptual dependency. Experimental results of speech understanding for various utterances concerning Kamakura sightseeing with a 1000-word vocabulary are also reported.

1 はじめに

我々は、自由度の高い発話表現を理解対象とした音声対話システムを目標に、概念素に基づく意味理解方式 [1, 2] の検討を進めている。本方式では、統語的な制約力を緩め、意味知識を主体とした制約力を言語尤度として適用することで言語的頑健さを実現している。しかし、文の意味的妥当さの観点では、概念素を意味的に関連付ける制約力が弱いため、意味をなさない対立候補を文意として受理していた。

この課題に対し、文献 [3] では言語尤度を算定するペナルティ規則 [4] の詳細化により、意味的曖昧さの検証を試みた。しかし、ペナルティ規則を詳細に設計することは、発見的知識やタスク知識の領域まで記述することになり、言語的制約知識の汎用性の観点から望ましくない。我々は今後、本方式の実用化にあたって、理解精度向上のためにタスク知識を積極的に利用することを考えており、ポータビリティの観点から、タスク知識を独立に扱う必要がある。

そこで、まず、本意味理解方式で用いている言語的制約知識からタスク知識を分離することを試み、言語的制約知識の一般化を行なった。次に、一般的な言語知識として導入した概念依存関係に基づいて、概念素間の部分依存構造を構築していく概念素の統合手法 [5] を検討した。本稿では、一般化された本意味理解方式を提案するとともに、語彙 1000 単語規模の鎌倉観光案内の自由な質問を対象とした音声理解実験を行ない、本手法の音声理解性能を評価した結果を報告する。

2 言語制約知識の一般化

本方式による意味理解動作の基本原理では、発話文中で意味的まとまりをなす複数の文節を中間的な理解単位である概念素として抽出し、これら概念素を統合して文意の意味表現を形成する。

従来は、概念素の種類をタスク分析により決定して意味フレームで表現し、文節を概念素フレームのスロットに入れる際に、文節の言語的特徴に

より概念素の言語的尤度を決め、概念素仮説を生成する。文意は、概念素をスロットとする意味フレームに概念素仮説を代入し、言語 / 音響尤度の高い意味フレーム表現を理解結果としていた。

このとき、言語的制約上の問題点として、(1) 概念素を構成する文節間の意味的依存関係を定義していないため、意味をなさない文節の組を受理していること、(2) (1) の条件下で制約力を強めるために、概念素の定義にタスク依存の知識を導入していたこと、(3) 概念素間の意味的依存関係を定義していないため、意味的に関連のない概念素の組を受理していること、などがあつた。そこで、基本的に次の二点に着目して言語的制約知識の一般化の検討を行なった。

・ (1) と (3) は、格関係や意味修飾関係に関する一般的な概念依存関係を導入することで、タスク依存性を除去すると共に、一般的な言語知識からみた妥当な意味関係を定義した。

・ (2) は、概念素を定義する言語制約知識は一般的な言語知識とし、タスク知識は、概念素から文意を構成する段階で、文の意味的妥当さの高い文意仮説を選択するタスクフィルタとして適用することとした。

2.1 概念素の定義

概念素の種類は定義では、設計者の恣意性をできるだけ避けるために、既存の概念分類体系の一つである文献 [6] を参考にした。これらの概念の中から鎌倉観光案内のタスクに出現し得るものを選定し、事物概念 (名詞)、動作概念 (動詞)、属性概念 (形容詞、副詞) に分類して、計 107 種類を設定した。また、概念間の階層化は、階層構造自体がタスク依存となる可能性があるため、行なわない。定義した概念素の例を、表 1 に示す。

2.2 概念依存関係知識の導入

概念素間の意味的依存関係について、

$\langle \text{概念 A} \rangle \Leftarrow (\text{依存関係子}) \langle \text{概念 B} \rangle$

の形式により表現する。例えば、「ホテルを探してるんですが。」という発話の場合、概念は「ホ

表 1: 概念素の定義例

事物：物、物資、乗り物、施設、人間、景観、自然(地勢)、状態(気配)、文書、行動(往来)、行動(約束)、位置(場所)、位置(周辺)、時(時間)、暦日(季節)、指示(場所)、指示(物)、不定(場所)、不定(物)。
動作：存在、授受、操作、労役、往来、出没、変動、加除、搜索、約束、思考、取引、行為(知覚)。
属性：性質(空間)、性質(形状)、性質(聴覚)、価値(価格)、価値(有用)、状態、関係(遠近)、時(終始)、程度。

テル」が<施設>、「探す」が<搜索>となり、両者を依存関係子 object により関係付ける。

<施設> <==(object) <搜索>

このような概念依存関係を、(1) 事物概念と動作概念との格依存関係、(2) 事物概念と属性概念(形容詞)との格依存関係、(3) 事物概念同士の意味的修飾関係、について設定する。(3)は、助詞「の」によって結ばれた名詞間の意味関係であり、依存関係子 modify により定義する。

3 概念素の統合処理

文節ラティスから概念素を抽出する際に、概念素間の依存関係を意味検証しつつ、概念依存関係の部分依存構造を構成するために、概念素の統合処理を行なう。

3.1 概念素の表現

概念素の表現は、図 1 に示すように、(1) 概念素名、(2) 依存関係で結ばれる概念素、(3) 概念素の属性、(4) 文節、の情報を持つ。(2)は、名詞や副詞と述語との格関係、及び名詞間の連体修飾関係で結ばれる概念素、(3)は、概念素の格性や、連体修飾性などの情報である。

3.2 概念素の統合動作

まず入力された文節ラティスから、各文節に対する初期の概念素仮説を生成する。同時に、概念素の属性も決定する。属性の決定には文節の形態素情報を用いるが、格助詞が省略された文節に対しては可能性のある格の仮説を立てて属性とする。

次に、各概念素仮説に対して、隣接する他の概念素のうち、概念依存関係を有する概念素を依存関係で統合し、二つの概念素間の依存構造を持つ新しい概念素仮説を生成する。この統合では、図 2 に示すように、概念素の中心的意味が概念素 A から概念素 B へ移ったと考え、新しい概念素仮説に概念素 B の名前と属性を伝搬させる。このような統合を繰り返して、連体修飾性の属性を持たない概念素仮説を受理し、文節ラティスから概念素間の部分依存構造のラティスを抽出する。

概念素の言語尤度は、概念依存関係を有する概念素間の概念共起確率で与えることとし、音声認識と統合する際には、音響尤度と言語尤度の総合尤度を用いる。

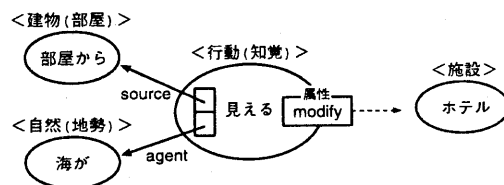


図 1: 概念素の表現を示す概念図

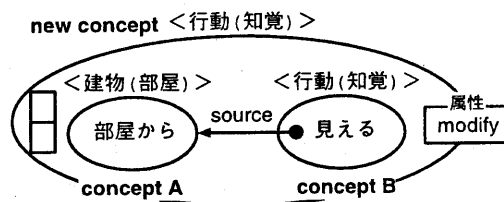


図 2: 概念素の統合動作を示す概念図

3.3 考察

以上の一般化された概念素統合方式に基づき、鎌倉観光案内タスクで概念素抽出性能の簡単な評価を行なった。入力とした文節ラティスは、50 評価文に対して、語彙 1379 単語の文節文法を用い、対立仮説 (False alarm) 数が入力文節数の約 20 倍程度となるように音声認識誤りを模擬して作成した。音響 / 言語尤度ともに用いず、全ての文節候補を探索して概念素抽出を試みた。

その結果、概念素仮説数の一文あたりの平均は62.0であった。同程度規模の文節ラティスに対し、従来の言語尤度の制約 [3] により生成された概念素仮説数の平均が、一文あたり数百程度であったことから、タスク知識を用いなくても、一般的な言語制約知識のみで十分に無意味な概念素仮説を棄却できることがわかった。

また、生成された概念素ラティスを詳細に調査した結果、意味的に妥当と考えられる概念素仮説の占める割合は正解も含めて58%であり、棄却可能な仮説が全体の約4割を占めていた。これらの棄却すべき仮説に関して検討したところ、以下のことがわかった。

まず、言語知識による制約の問題としては、概念依存関係は格関係で結ばれた二つの概念を制約するものであり、格要素となった概念素同士を制約しないことが挙げられる。例えば、「九時から(時-起点格)、あるんですか」や、「庭(主格)、あるんですか」などは意味的に妥当だが、「九時から、庭(主格)、あるんですか」の例は妥当でない。これは、存在を表わす動詞「ある」が非常に基本的な動詞で多くの意味を持つことから生じる。このような場合は、タスク知識を適用する以前に一般的言語知識により制約が可能で、動詞の多義に対してそれぞれの動作概念を定義しておく方法が考えられる。

次に、概念素の統合の問題では、連体修飾関係の意味制約が弱い。連体修飾性の動作概念を後続の事物概念に統合する際、後続の事物概念が動作概念の格関係のいずれかに適合すれば連体修飾が成立するとしていた。しかし、「食べられる(連体形)、お寺」といった、主格を伴わない連体修飾の例があった。これは、「食べる」が「お寺」を場所格として統合した結果であり、「精進料理が(主格)、食べられるお寺」のように主格を伴えば全く妥当な理解結果である。さらに、省略発話や断片的な発話を理解対象とするためには、このような概念素仮説を棄却することはできない。そこで、言語的頑健さの観点から、一般的な言語制約では受理、または優先度付けに留めておき、文

脈まで考慮した後段のタスクフィルタにより最終的な理解結果を決定する方法が良いと考える。

4 概念素から文意表現への構成

概念素ラティス中の概念素仮説を統合して、文意表現を構成する。

従来 [1] では、文意を発話意図と概念素との組であると定義し、発話意図は、例えばホテル予約のタスクに関して、「Yes-No 質問、各 WH 質問、予約、検索、変更、取消」などのように分類した。これらには、表層的な表現で決定すべきもの (Yes-No 質問、各 WH 質問) と、述語の意味的な内容を表現するもの (予約、検索、変更、取消) とが含まれており、後者はタスク依存であった。

そこで、発話意図の定義を一般的なものとするために、発話意図を、(1) 「真偽質問、WH 質問、要求、希望、可能、宣言」といったモダリティ情報と、(2) 動詞や形容詞などの叙述的概念とに分離して、両者の組により表現する。さらに、発話意図の叙述的概念による意味的制約を用いるために、各叙述的概念が持つ格構造(文型)を概念依存関係により定義する。従って、一般化した文意の定義は以下ようになる。

<文意> :=

<発話意図(叙述的概念+モダリティ)>

+ <叙述的概念に統合される概念素の組>

文意表現は各叙述的概念に対応して定義し、以下の構成要素を格納する。

- 叙述的概念：モダリティ情報に関する属性を持つ概念素
- モダリティ情報：叙述的概念素により決定
- 統合概念素の組：属性が格性、連用修飾性、連用節性の概念素

叙述的概念が省略された発話の場合でも、各文意表現に対応する叙述的概念の概念依存関係を用いて、概念素仮説を統合する。

このような文意表現から構成される意味表現は、一般的な概念依存関係により関連付けられたもの

であり、タスクに依存しない。タスク知識を利用した制約は、一般的な文意表現から、システムの入力となるタスク依存の意味表現への意味翻訳過程で行なうことを基本とする。ただし、タスク知識による制約は、概念素の抽出と統合、及び、文意表現の構成の過程へフィードバック可能であると考えている。

5 音声認識との接続手法

本意味理解手法を音声認識と組み合わせるために、(1) 概念素の統合処理において、文節ラティスからの島駆動探索による概念素抽出手法 [1, 7]、及び、(2) 文意表現の構成処理において、効率的な文意仮説探索手法 [1, 8] を用いる。この文意仮説探索手法は、理解処理の早期で正解概念素の脱落を許容し、信頼度が高い概念素仮説からなる初期文意仮説を高速に得て、この仮説から想起される概念素制約知識をトップダウンに利用しながら、脱落区間を再探索して正解の概念素仮説を救う手法である。ここでは、概念素制約知識として、文意表現の叙述的概念に対応して定義されている概念依存関係を用いる。

6 音声理解実験

一般化した本意味理解方式の理解性能を評価するために、音声理解実験を行なった。

6.1 実験条件

評価タスクは、鎌倉観光案内システムを構築することを想定し、(1) 観光名所案内、(2) 飲食店案内、(3) ホテル案内・予約、(4) 道案内、を含むタスクである。評価音声は、被験者が実際の対話を想定して自由に作成した 50 質問文を、男性 3 名が読み上げた発話データを用いた。

音声認識の文節スポッティングは、文献 [9] の未知語処理なしの方式を用いた。文節内ネットワークは不特定話者の音節 HMM で構成され、ネットワークの規模は語彙 970 単語 (自立語 836、付属語 134) である。なお、この文節ネットワークの語彙が、考察時に使用した文節文法の語彙と異なる

理由は、非文の過剰生成を抑えるために文節文法を再検討したためである。スポッティングで得られた文節ラティスの精度を表 2 に示す。

意味理解では、概念素数 99、概念依存関係は 13 種類、文意表現は 35 種類を用いた。文節候補の時間的接続の許容範囲は、overlap と gap の両者とも 12 フレーム (フレーム周期 10 ms) とした。概念素仮説と文意仮説の尤度は、スポッティングの音響尤度のみ用いた。理解結果の出力は一般的な文意表現による意味表現であり、タスク知識は用いていない。文意の理解結果に対する正解の認定基準は、(1) 発話意図 (叙述的概念+モダリティ)、(2) 各概念素の種類と境界、(3) 概念依存関係、(4) 文節の自立語、の全てが正しいことを条件とした。

表 2: 評価文節ラティスの文節検出率 (%) : 50 評価文。語彙 970 単語。規模は、入力文節数に対する false alarm 数の割合 (倍) を示す。

規模	話者 A	話者 B	話者 C
1.2	78.0	79.9	64.5
4.9	93.0	98.1	90.7
8.9	96.3	98.6	94.9
18.9	97.7	99.1	97.7
38.9	98.1	100.0	98.6

6.2 実験結果

文意の理解率を表 3 に示す。文節ラティスの精度が比較的良好な話者 A と B では、第 1 位で 86 ~ 88%、第 10 位までの累積で、ほぼ全ての正解を得ており、良好な理解結果を得た。しかし、ラティス精度が劣る話者 C では第 1 位で 70% であり、良好な理解精度とは言えない。話者 C に関して理解結果を視察により分析した結果、以下のことがわかった。

まず全般的に、理解誤りの文意候補は、タスク知識で棄却可能な概念素を含むものが多く、タスク知識導入により理解精度を向上させることができる (例: 「かど (飲食店名) のツイン」、「うどんの入場料」、「空いてる時間、知って下さい」)。

次に、言語的な要因の理解誤りでは、考察で議論したように、概念依存関係で結合された二つの概念素の意味共起が妥当であっても、文意の格要素となる概念素同士の意味共起が適切でないものが多くみられた(例:「鎌倉パークホテルってどこにあるんですか」→誤り「鎌倉パークホテルってどこにやるんですか」)。

一方、音響的な要因の誤りでは、格を特徴付ける上で重要な助詞(「から」、「まで」など)の誤りに起因する概念素の属性誤りがあった(例:「近代美術館って何時まで空いてるんですか」→誤り「近代美術館って何時、まだ(副詞)空いてるんですか」)。この例の場合は、助詞の誤りが生じていても重要な自立語は正しく検出できているものの、語彙をさらに拡張した場合には、この属性誤りは増加すると予想され、助詞誤りを考慮した概念素属性の決定が課題である。また、同時に助詞検出精度の高い音声認識も課題となる。

最後に探索の観点からは、話者Cの場合に再探索に失敗した例が3例あった。脱落を許容する初期文意仮説の精度を向上させる必要がある。

表 3: 意味理解率(%): 50 評価文。語彙 970 単語。Rank は文意候補の順位を示し、第 2 位以下は累積理解率を示す。

Rank	話者 A	話者 B	話者 C
1	88	86	70
2	92	88	78
3	94	90	82
5	96	94	82
10	98	100	86

7 おわりに

言語制約知識を一般的な概念依存関係により表現し、概念素間の部分依存構造を構築していく概念素理解方式を提案した。音声理解実験による評価の結果、語彙 1000 単語規模のタスクで、良好な理解性能を得る可能性を確認した。今後は、複数の概念依存関係にわたる広範囲な意味制約、及

び、タスク知識による意味制約手法の検討を行ない、理解精度向上を図る。

謝辞: 研究の機会を与えて下さった、三菱電機(株)情報総研 音声・言語インタフェース技術部 部長 中島邦男氏 と、音声認識の実験データを提供して下さいました、同 花沢利行氏に感謝致します。

参考文献

- [1] 永井明人, 石川泰, 中島邦男 “概念素に基づく意味理解における音声認識への統合手法,” 情報処理学会 音声言語情報処理研究会, SIG-SLP 7-4, pp. 23-28 (Jul. 1995).
- [2] A. Nagai, Y. Ishikawa and K. Nakajima, “Integration of Concept-Driven Semantic Interpretation with Speech Recognition,” Proc. ICASSP’96, Atlanta (U.S.A.), pp. 431-434, (May 1996).
- [3] 永井明人, 石川泰, 中島邦男 “概念素に基づく意味理解における言語知識の詳細化,” 日本音響学会 春季研究発表会 講演論文集, pp. 173-174 (Mar. 1996).
- [4] 永井明人, 石川泰, 中島邦男 “概念素に基づく音声対話理解方式における意味スコアリング法,” 日本音響学会 春季研究発表会 講演論文集, pp. 151-152 (Mar. 1994).
- [5] 永井明人, 石川泰, 中島邦男 “概念素理解方式の一般化とタスク知識の適用手法,” 日本音響学会 春季研究発表会 講演論文集, pp. 45-46 (Sep. 1996).
- [6] 大野晋, 浜西正人 「角川類語新辞典」, 角川書店.
- [7] 永井明人, 石川泰, 中島邦男, “文節ラティスからの概念素探索手法,” 日本音響学会 秋季研究発表会 講演論文集, pp. 35-36 (Oct. 1994).
- [8] 永井明人, 石川泰, 中島邦男, “概念素探索法における概念素仮説からの文意統合方式,” 日本音響学会 春季研究発表会 講演論文集, pp. 69-70 (Mar. 1995).
- [9] 花沢利行, 阿部芳春, 中島邦男, “文節スポッティングにおける未知語検出方式の改良,” 日本音響学会 秋季研究発表会 講演論文集, pp. 179-180 (Sep. 1996).