

| | |
|--------------|---|
| Title | 機械可読辞書中の語釈文を用いた動詞間の時間関係の解析 |
| Author(s) | 森部, 正二郎 |
| Citation | |
| Issue Date | 1997-03 |
| Type | Thesis or Dissertation |
| Text version | author |
| URL | http://hdl.handle.net/10119/1040 |
| Rights | |
| Description | Supervisor:奥村 学, 情報科学研究科, 修士 |

機械可読辞書中の語釈文を用いた 動詞間の時間関係の解析

森部 正二郎

北陸先端科学技術大学院大学 情報科学研究科

1997年2月14日

キーワード: 機械可読辞書, 時間関係, 一般知識, 語義ネットワーク.

テキストの抄録作成など、テキストの知的処理ではテキストの内容を理解することが重要な役割を果たす。テキスト理解には、テキスト中の動詞間の時間関係を明らかにすることが必要であるが、通常のテキストは必ずしもイベントの発生順に記述されているとは限らない。

本研究では動詞間の時間関係の解析を目的とする。動詞間の時間関係を解析する方法には

手法 1. テンス・アスペクト・時間指示語などの時制情報を利用する

手法 2. イベント間の時間関係に関する一般知識を利用する

などの方法がある。[手法 1] は、複文中の動詞間の時間関係を解析する場合に用いられるが、時制情報だけでは曖昧さが生じることが多い。[手法 2] を用いることの利点は、時制情報に関係なく時間関係を決定できることである。従って、[手法 1] による曖昧さの解消や、2つの単文における動詞間の時間関係も解析することができる。2つの単文の時制が一致していて、「その後」、「その前に」のような時間指示語がない場合、[手法 1] で時間関係は解析できない。

東らは時間に関する言語情報と一般知識や記述分野の専門知識の両方を用いている [東 96]。さらにそれらの知識を統合する Event Calculus の枠組を使ってイベント間の相対的な時間関係を抽出している。特徴は、一般知識と専門知識を分離しているため、専門知識を変えるだけで様々な分野の文章で扱えるという点である。また Alterman は知識として動詞間の時間関係を意味ネットワークとして用いている [Alterman 85]。この研究ではその意味ネットワークから、入力テキスト中のイベント間の関係だけでなく、各イベント

に関連して生じるイベントとその関係も出力する。東条は、一般知識を用意する代わりに各イベントに時間的特性を与える方法を提案している [東条 95]。イベント間の関係は時間特性に基づいて定義された時間区間どうしの関係の付け方によって決定される。

これらの研究と我々の研究の違う点は、一般知識や専門知識を知識ベースとしてあらかじめ人手で作成しないという点である。人手で作成する場合、手間がかかる上、客観性を欠いた知識になる危険性がある。我々はこれまで、時制表現を考慮した解析に加え、動詞間の依存関係を導くために発生や消滅などの状態変化を表す動詞に注目してきた [森部 95]。この手法は動詞の分類と状態変化する対象を調べるだけで依存関係がわかるが、この条件を満足する文は非常に少ない。本研究ではより多くの依存関係に対応するために、一般知識として計算機用日本語基本動詞辞書 IPAL [情処 87] に記述されている語釈文を情報源として用いる。

本研究では、イベント間の依存関係や順序関係に関する知識を計算機用日本語基本動詞辞書 IPAL の語釈文から半自動的に導く。語釈文は、

- 動詞を別の単語で言い替えたイベント・状態 (文の主動詞)
- 動詞の原因、目的、方法などを表すイベント (文の従属節)
- 動作主や対象の状態を表す格要素

から構成されている。語釈文からこれらのイベントや状態を抽出し、時間関係でリンクした意味ネットワークを、動詞の語義が表す語義ネットワークとする。この語義ネットワークを全ての動詞に対して作成し、一般知識として蓄積する。

時間関係解析システムは、2つの単文を入力とし、語義ネットワークを用いて単文中の動詞間の時間関係を出力する。動詞間の時間関係はこのシステムを利用して、次のように解析される。まず単文中の動詞の語義ネットワークを作成した一般知識から参照し、次にそれぞれの語義ネットワーク中に一致するイベントや状態が存在するか調べる。一致するイベントや状態の時間関係を同時とみなすことにより、単文中の2つの動詞が時間関係のリンクで結ばれる。結ばれた時間関係のリンクをたどって、単文中の2つの動詞間の時間関係を計算する。

論文では、まず語釈文から語義ネットワークを自動的に生成する手法を示し、次に語釈文から抽出した一般知識を用いて動詞間の時間関係を導く方法を示す。

日本語基本動詞辞書 IPAL の語釈文を形式化することにより、見出し語の意味を表す語義ネットワークを作成した。その知識を利用して、依存関係のある動詞の時間関係を解析する手法を示した。実際に時間関係を解析するシステムを実装し、新聞記事、物語文を対象に時間関係を解析した。結果として、新聞記事では約1割、物語文では6割の時間関係が導かれた。新聞記事では再現率は低いですが、時制情報のみを用いて時間関係を解析する方法で解析できない部分の時間関係をカバーしていた。