

論文要約

学位論文題目：「日本語談話関係認識のための理論とコーパス構築」

氏名：金子 貴美

自然言語処理研究の究極的な目標の1つは「計算機に文章・文脈を理解させること」である。文章は、それを構成する各文を意味解析しただけでは導出できない、文間の構造に関わる意味を持っているため、上記目標を実現するには、文間や節間にある意味的な繋がり関係を解析する必要がある。このような、文や節といった談話片間の意味的關係のことを談話関係と呼ぶ。例えば、以下の2つの文章を考える：

- (1) a. 花子は落とし穴に落ちた。
b. 太郎は花子を助けた。
- (2) a. 花子は落とし穴に落ちた。
b. 太郎が花子を押し込んだ。

(1)(2)に示す例における2文間の時間関係は異なっている。(1)では、「落ちた」という出来事と「助けた」という出来事が起きた順序は文の順序と合っている一方、(2)では、「押し込んだ」ことにより「落ちた」と解釈できるため、出来事の成立順序と文の順序は合っていない。文の順序に沿って、文の意味解析結果を繋ぎ合わせて文章の解析結果とした場合、(1)(2)の時間配置を区別することができない。しかしながら、談話関係が解析できれば、これらを区別することができる。このように談話関係の理解は、文章内の出来事間の時間的・空間的(spacio-temporal)な関係を捉える上で重要な役割を果たすため、自然言語理解に必須のモジュールであると考えられる。

一方、談話関係は(1)(2)のように談話関係が決まると時間関係が特定されるものだけでなく、時間関係が決まると談話関係が特定されるものがある。次の2つの文章について考える：

- (1) a. カーテンを開けた。
b. 一面の銀世界が広がっていた。
- (2) a. a. カーテンを開けた。
b. 一面の銀世界を見下ろした。

(3)(4)に示す例における2文間の時間関係も異なっている。(3)では、「開ける」動作の成立時に「広がっている」状態が成立していたと解釈できる。一方(4)では、「開ける」の動作の後、「見下ろす」動作を行ったと読み取れる。このように、談話関係と時間関係は相互に依存しているため、そのような依存関係を考慮して談話関係の定義がなされるべきである。

しかしながら、談話関係に関するいずれの先行研究も、話関係と時間関係の依存関係を反映した設計となっているとは言い難い。そのような設計ができていない一因として、いずれの先行研究も解釈が揺れる「話し手や書き手の意図に関する関係」と、解釈が一意に定まる関係とを十分に選り分けられていないことがある、と推察される。

上記の問題点を踏まえ、本研究では、話し手や書き手の意図に関する関係と、解釈が一意に定まる関係とを選び分け、後者について、時間関係との依存関係を反映した、談話関係の決定手順を確立することを目指し、談話関係を再定義する。具体的には、上記の実現のためにまず、出来事 (eventuality)、モダリティ (modality)、時制 (tense)、時間関係 (temporal relation) と、談話関係との依存関係の整理を行い、それを元に談話関係の分類体系と客観的な談話関係の決定手順を定義する。その後、定義した分類体系と決定手順に従ってアノテーション実験を行い、一致率の算出やエラー分析を行うことにより、「解釈が一意に定まる関係のみに絞って談話関係を定義できたか」を検証し、談話関係の決定プロセスのどの部分を明らかにできたか、および、どの部分に課題が残るかを議論する。