

日本語文のテンス・アスペクト構造解析

谷川 雄亮 横山 晶一 西原 典孝
山形大学大学院 理工学研究科

1. はじめに

われわれが、言語を用いて対話を行う際、そこには必ず時間的な概念を伴う。したがって対話において時間関係を把握することは重要である。しかし日本語には英語のように厳密に定められた時制が存在せず、日本語の時間表現はあいまいであると言われる。

前研究[1]では動詞の種類、アスペクトに着目し、文中におけるそれらの組み合わせから基本的な時制を機械的に捉えるアルゴリズムを構築した。

本研究では前研究の成果を踏まえた上で、時を表す副詞的表現、アスペクトについてさらに深く追求し、その文が表す時間的な局面、時間構造をより明確に捉えることのできるアルゴリズムを構築した[2]。その結果、日本語文の持つ基本的なテンス・アスペクトの意味的構造を捉えることが可能となった。

2. 背景

2.1 テンスとアスペクト [3]

歴史的な時間の流れの中で、事象がいつ起こるか、いつ起こったかを示す文法体系をテンスといい、事象を示す動詞が表す動作や状態の時間的な局面、例えば開始（～し始める）、終了（～し終わる）などを表す文法体系をアスペクトという。

日本語のテンス・アスペクト体系は「スル、シタ、シテイル、シテイタ」という4つ語彙の対立として扱われている。

2.2 動詞の分類

テンス・アスペクトは、動詞の語彙的内容とも深く関わっているため、動詞の分類が必要となる。本研究では動詞を時間的な性質から分類した工藤[3]の分類を採用し、次の3種に分ける。

A. 外的運動動詞：物理的な動作

(A.1) 主体動作・客体変化動詞：開ける

(A.2) 主体変化動詞：集まる、開く

(A.3) 主体動作動詞：走る、歩く

B. 内的情態動詞：

感情や思考など、人の内的な事象：思う、信じる

C. 静態動詞：

状態、特性などの静的な事象：ある、存在する

2.3 時の副詞 [4]

時間表現を捉える際に有効な時を表す副詞的表現として、基準となる時間(主に発話時)との相対的な関係を表す<時の状況成分>と、時間の中で事象の出現、展開、存在のありようなど事象の内的な時間的特性を表した<時間関係の副詞>がある。

・ 時の状況成分の例

今日、今週(発話時を含む時間帯)

昨日、この前(発話時以前)

明日、来週(発話時以後)

1999年、5月5日(絶対的時点)

現在、目下(発話時点そのもの)

・ 時間関係の副詞の例

ずっと、しばらく(事象存続の時間量)

だんだん、徐々に(事象の進展)

すぐ、やがて(起動への時間量)

いつも、ときどき(頻度)

毎日、毎年(繰り返し期間)

2.4 基本的時制の判定

前研究[1]において動詞のアスペクト、動詞の組み合わせから基本的な時制の判定を行うことに成功している。ここではその一例をあげる。

- ・彼は50メートルを泳ぐ。
<動作動詞+スル形> 現在または未来の事象

3. 日本語のテンス・アスペクト[5]

本研究では、日本語のテンス・アスペクトについて以下のものを扱う。

完成相：

出来事の始めから終わりまでをひとまとまりとして捉える

- ・花子はカレーを作った。(完成相過去)

継続相：

出来事の進展の中での一段階を取り出して示す

- ・太郎は本を読んでいる。(動作継続相現在)
- ・窓が開いている。(結果継続相現在)

パーフェクト相：

発話時以前に起動し既に終了している事象が現在においても有効、または既成の事実として記録

- ・太郎は5年前に死んでいる。
(パーフェクト相現在)

反復相：

同じ出来事が反復的に生起することを表現

- ・父は毎日晚酌をする。(反復相現在)

アスペクト性を持たない静的事象：

- ・その引き出しの中に書類がある。
(アスペクト性を持たない状態現在)

4. 動詞のアスペクト

本研究ではスル、シタ、シテイル、シテイタという基本形のアスペクトの他に以下のものについても扱う。ここであげる動詞のアスペクトは、それだけで文全体のテンス・アスペクト構造を決定できることが多い。

- シテアル、シテアッタ (結果継続相)
- シハジメル、シハジメタ (開始相)
- シオウル、シオワッタ (終了相)
- スルコトガアル、スルコトガアッタ (反復相)
- シタコトガアル、シタコトガアッタ
(パーフェクト相)

5. 日本語文のテンス・アスペクト 構造解析アルゴリズム

5.1 解析の流れ

機械処理によって、入力された日本語文のテンス・アスペクト構造の解析を行う際の大まかな流れは以下ようになる。

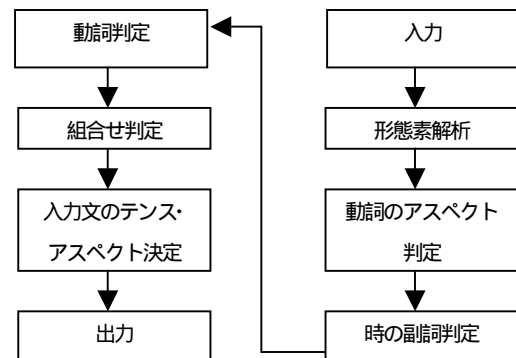


図1 テンス・アスペクト構造解析の流れ

5.2 テンス・アスペクト 構造決定要素の優先順位

アスペクト、時を表す副詞的表現、動詞の各要素について日本語文のテンス・アスペクトを決定する際の優先順位を影響の強さによって設定した。

- 頻度副詞・繰り返し期間を表すもの
- スル・シタ・シテイル・シテイタ以外の特殊型アスペクト
- 静態動詞
- 起動への時間量
- 事象の進展

5.3 日本語文のテンス・アスペクト 構造解析アルゴリズム

これらの内容をもとに構築した、スル形アスペクトの日本語文のテンス・アスペクト構造解析アルゴリズムを図2に示す。このアルゴリズムでは、入力された文を形態素解析し、前節のテンス・アスペクト決定要素を含んでいるかどうかを優先順位にしたがって判定してアスペクトを決定する。その後、時の状況成分によって最終的にテンスを決定し出力を得る。

シテイル形アスペクト文の場合はスル形での処理にパーフェクト相、動作継続相、結果継続相の判定が加わるため、時の状況成分と動詞の種類の判別が重要になる。

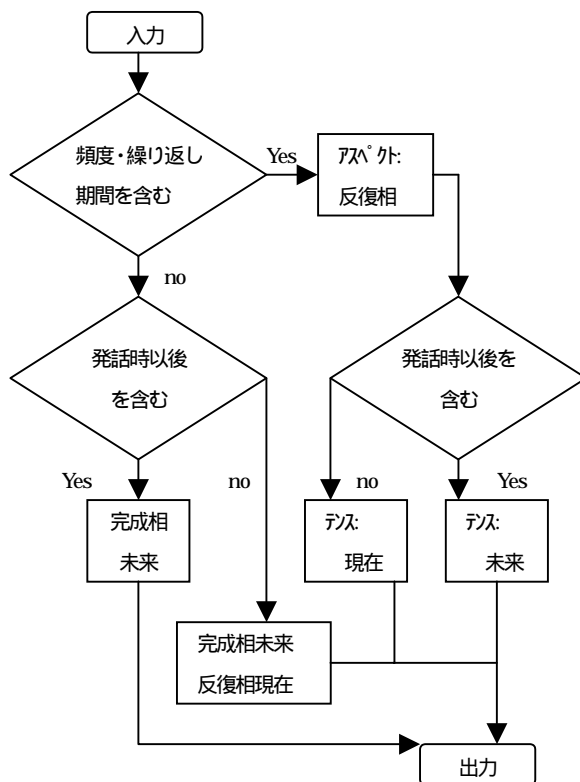


図2 スル形アスペクト文の
テンス・アスペクト構造解析アルゴリズム

以下にこのアルゴリズムに従って解析を行った例を示す。

私は、毎日歯を磨く。

毎日 繰り返し期間
磨く 主体動作・客体変化動詞、スル形
以上の条件から「反復相現在」

彼は、明日ここへ来る。

明日 発話時以後
来る 主体変化動詞、スル形アスペクト
以上の条件から「完成相未来」

彼は服を着ている。

着る 再帰動詞
ている シテイル形アスペクト

文中に含まれるテンス・アスペクト決定要素が再帰動詞のシテイル形アスペクトのみなので、動作継続と結果継続の最終的判定ができない。また、反復相・パーフェクト相の可能性も残る。

6. 考察

6.1 本アルゴリズムの精度

E D R日本語コーパス辞書[6]から任意に抽出した400文について、本アルゴリズムに従って、人手でテンス・アスペクト構造の解析を行ったところ、最終的に1つのテンス・アスペクト構造を捉えることができ、なおかつそれが正しい判定と言えるものを正解として、約84.5%の正解率を得た(400文中338文)。詳細は下記の表1に示すとおりである。しかし、本アルゴリズムでは正しい解析結果を得られない文や、文中の情報不足により複数の解析結果が出力されてしまう例もあるため、それらへの対処法を検討する必要がある。また、日本語の表現は多彩であり、このアルゴリズムが適用できない表現方法が存在することも考えられるため、さらに多くの例文について解析を行う必要がある。

表1. アスペクト毎の正解率

	正解数	正解を含む複数	エラー	一意正解率
完成相	68	6	0	91.1%
動作 継続相	65	12	3	76.9%
結果 継続相	63	13	4	73.0%
反復相	60	10	3	78.3%
パーフェクト相	52	8	3	78.8%
その他	30	0	0	100%

6.2 誤判定の例

このアルゴリズムでは、以下のように正しい解析結果が得られない場合がある。

- ・彼は、食べ物の好みが相当変わっている。

このような例では、「変わる」という主体変化動詞で表されている事象であり、本アルゴリズムを適用すると「結果継続相現在」と判定される。しかし、この文が実際に表している意味内容は主体の特性を表す、「アスペクト性を持たない静的状態」である。

6.3 複数の出力を得る例

文中の情報が不足すると、次のように可能性のあるものを全て出力することになる。

- ・次郎は今、服を着ている。
 - ・動作継続相 服を着ている最中
 - ・結果継続相 服を着た状態

このように、再帰動詞や二側面動詞などの動作と変化の両方の局面を持つ動詞で示される事象は、動作継続相と結果継続相の判定を行うのが非常に困難である。ただし、次のような条件が加わることで最終的なアスペクトの決定が可能であると考えられる。

- ・次郎は今、赤い服を着ている。
(対象物の特徴 結果継続)
- ・次郎は今、あわてて服を着ている。
(動作を修飾するもの 動作継続)

7. 結び

本研究では、厳密に定義された時制が存在せず、あいまいであると言われる日本語の時間表現について、文中の動詞の種類、動詞のアスペクト、時の副詞という各要素から基本的なテンス・アスペクト構造を捉えることが可能となった。しかし、現段階では多彩な日本語の表現に対応するには不十分であり、前節であげたような問題点も残っている。今後、このような問題点を解決し、さらに精度の高い解析を行うことができるアルゴリズム・システムを構築することが今後の課題となる。

また、翻訳システムや対話システムをはじめとした、自然言語処理システムにおける時間認識への応用の検討も必要である。

参考文献

- [1] 谷川雄亮：アスペクトによる日本語時間表現の分類、山形大学 卒業論文 (2002)
- [2] 谷川雄亮：日本語文のテンス・アスペクト構造解析、山形大学 修士論文 (2004)
- [3] 工藤真由美：アスペクト・テンス体系とテキスト、ひつじ書房 (1995)
- [4] 仁田義雄：副詞的表現の諸相、くろしお出版 (2002)
- [5] 金水敏、工藤真由美、沼田善子：時・否定と取り立て、岩波書店 (2000)
- [6] EDR日本語電子化コーパス、日本電子化辞書研究所