

TR-IT-0221

デフォルト格ラベル自動付与実験

Automatic Tagging of Default Case Labels

- An Experiment -

荒川直哉

ARAKAWA Naoya

1997年6月

概要

This is a report on an experiment of automatic case analysis in Japanese. Dependency structures of the ATR travel domain corpus (SLDB) are automatically tagged with case labels. The rules for tagging were formed mostly with the information from part of SLDB which has been tagged with case labels by hand. The report presents the result of the experiment and discusses problems.

© (株)ATR音声翻訳通信研究所 1997

© 1997 by ATR Interpreting Telecommunications Research Laboratories

# 目次

1 概要 .....	1
2 デフォルト格ラベル自動付与規則 .....	1
2.1 規則の種類と適用順序.....	1
2.2 TGRAM (三項規則).....	2
2.3 BGRAM (二項規則).....	3
2.4 PAT.....	5
2.5 PAT2.....	7
2.6 EMPTY.....	7
2.7 DEL.....	8
2.8 DEL2.....	9
2.9 ソーティング .....	10
2.10 REL .....	10
3 助詞類の出現パターンによる処理方法 .....	10
3.1 助詞の連続.....	10
3.2 助詞の交替.....	12
3.2.1 コントロール現象による助詞の交替.....	12
3.2.2 強調による助詞の交替.....	13
3.3 助詞の省略.....	14
3.4 <副助詞>の扱い.....	14
4 実験結果.....	15
4.1 評価の概要.....	15
4.2 評価の方法.....	15
4.3 評価結果.....	15
5 問題点と改良案.....	18
5.1 関係節の場合の格の二重付与 .....	18
5.2 DEL による副作用.....	19
5.3 EMPTY に登録されている省略格パターンの整備.....	20
5.4 データサイズとスパース性.....	22
5.4.1 見出しの一部を用いる方法.....	22
5.4.2 意味情報を用いる方法.....	23
参考文献.....	24
付録 .....	25
1 深層格のリスト.....	25
2 ファイルリスト.....	26

## 1 概要

このレポートは、格構造解析における深層格ラベル付与を自動的に行う実験についての報告である。ここで格構造解析とは、依存構造解析<sup>1</sup> 結果中の用言の各格要素に対して深層格を明示する情報（深層格ラベル）を与える作業を指す。深層格ラベルは述語と格要素の意味的な関連を表示する。したがって、深層格付与は意味解析処理の一部をなし、一種の多義性解決処理である。深層格ラベルには、助詞（格助詞、係助詞、引用助詞）でマークされた要素に与えられるもの、関係節に与えられるもの、さらに表層形のない「省略格」（いわゆるゼロ代名詞）に与えられるものがある。本実験において、格助詞、係助詞への格付与、深層格の認定、関係節への格付与は、すでに深層格ラベル付与が行われたコーパス（文献 [2]）から情報を抽出して、格付与プログラム用の規則を自動的に生成することにより行われた。なお、若干の特殊ケースに対しては人手でデフォルト格ラベル付与規則を作成した。

本実験の目的は、

- 1) 用言とその格要素の表層情報を用いた深層格ラベル決定の精度<sup>2</sup>を調べる
  - 2) 人手による格構造解析の効率を高める
- ことである。今回の実験はあくまで予備的なもので、学術的な基準を満たすような網羅的なものではない。

実験に使用したコーパスは、ATR音声翻訳通信研究所で作成した旅行会話音声言語データベース（SLDB；文献 [1, 2]）の一部である（コーパスファイルのリストを付録2に付す）。実験で付与対象となった格構造データファイル用深層格ラベルの一覧を付録1に付す。

## 2 デフォルト格ラベル自動付与規則

### 2.1 規則の種類と適用順序

概要で述べたようにデフォルト格ラベル自動付与は、すでに格付与が行われたデータファイルから情報を抽出して、自動付与プログラム用の規則を自動的に生成することにより行われた。デフォルト格ラベル自動付与用の規則として次の種類を用意した。

- ・三項規則（TGRAM）

用言<sup>3</sup>、格（係）助詞、さらに格（係）助詞に依存する要素の組合せから、最も尤もらしい深層格ラベルを格（係）助詞ノードに付与する規則

- ・二項規則（BGRAM）

用言と格（係）助詞の組合せから、最も尤もらしい深層格ラベルを格（係）助詞ノードに付与する規則

- ・〈助動詞〉「だ」「です」用 [IDEN] 格付与規則（PAT<sup>4</sup>）

- ・〈引用助詞〉への [CONT] 格付与規則（PAT2）

<sup>1</sup> 依存構造解析とは、構文解析結果を主辞・従部関係を明示するような構造へ書き換える作業を指す。旅行会話音声言語データの依存構造解析結果は素性構造の形式を持つ（文献 [2]）。

<sup>2</sup> 「4 実験結果」を参照のこと。

<sup>3</sup> および体言につく助動詞「だ/です」。次の BGRAM についても同様。

<sup>4</sup> 元々はパッチ規則の意。なお、2.4 で説明するように、準体助詞「の」には [OBJE] が与えられる。

- 省略格付与規則 (EMPTY)

必須と思われる深層格が TGRAM, BGRAM, PAT, PAT2 を経ても付与されていない場合に、その格ラベルを追加する規則

- 余剰省略格削除規則 (DEL)

助動詞「だ」「です」に付与された余分な [OBJE NIL] を削除する規則

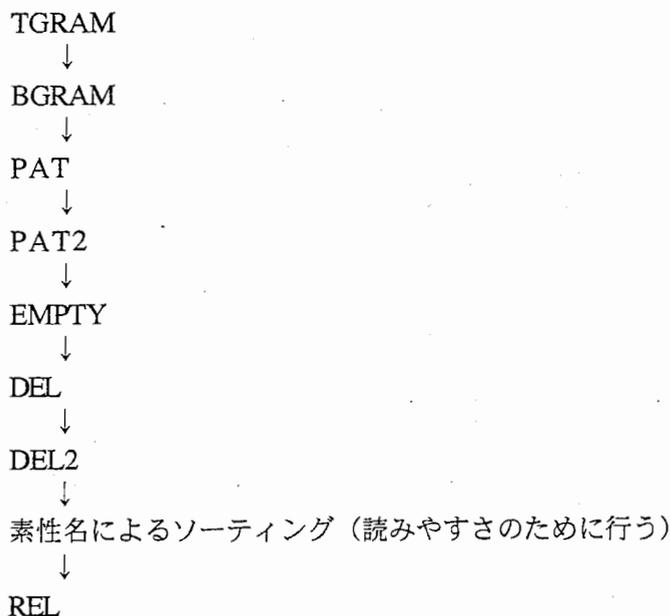
- 余剰省略格削除規則 2 (DEL2)

対応する深層格ラベルが係助詞を介して現われている場合に省略格ラベルを削除する規則

- 関係節格付与規則 (REL)

体言を修飾する述語に対して深層格関係ラベルを与える規則

規則の適用順序は以下の通りである。



ここで、TGRAM から DEL2 までの規則は、素性構造書き換えプログラム<sup>5</sup> の規則であるが、REL は専用のプログラム (2.10 参照) のための規則である。また、TGRAM, BGRAM, EMPTY, REL は統計に基づいて自動生成されたが、PAT, PAT2, DEL, DEL2 は人手により作成された。DEL, DEL2 は、規則の制御上の問題に対処するために導入されたもので、書き換え規則の記述能力が向上すれば不必要になる。

以下、各段階の操作と使用する規則について順に説明を行う。

## 2.2 TGRAM (三項規則)

TGRAM は、用言 (または <助動詞>「だ/です」)、格 (係) 助詞、さらに格 (係) 助詞に依

<sup>5</sup> ATR 音声翻訳通信研究所第 4 研究室で開発された汎用素性構造書き換えプログラムである。以下の本文中の例に現われるような書き換え規則を読み込み、入力される拡張された素性構造の中で規則に合致する部分を書き換えて出力する。

存する要素の組合せから、最も尤もらしい深層格ラベルを格（係）助詞ノードに付与する規則群である。ここで用言とは <本動詞>、<形容詞>、<形容名詞> および動詞用法の <サ変名詞>（依存構造データファイルでは [動詞 +] のマーキングがある）を指す。TGRAM は、手作業で格ラベル付けを行った格構造データファイルに関して上記の三項関係の統計を取り、その統計に基づいて作成される。

三項関係の統計例：

```
[[見出し      あ]
 [品詞        <本動詞>]
 [活用タイプ  五段ラ]
 [助詞        に]
 [目的語見出し 前]
 [目的語品詞  <普通名詞>]
 [生起回数    4]
 [統計リスト  (:LIST (LOCT 3) (TLOC 1))]]
```

この例は、統計を取ったコーパス中に「前に...あ(る)」という表現が4回生起し、そのうち「前」が場所 (LOCT) を表わす場合が3回、時間 (TLOC) を表わす場合が1回であることを示している。

TGRAM 規則例：

上記の統計例から生成される規則。統計では TLOC より LOCT の方が頻度が高いので、「前に...あ(る)」という表現中の格助詞「に」に対しては LOCT を与えることにする。

```
(defrule
:IN [[ヘッド !x01 [[見出し      あ]
                    [品詞        <本動詞>]
                    [活用タイプ  五段ラ]
                    ?rest0]]
 [引数   !x02 [[ヘッド [[見出し に]
                        [品詞    <格助詞>]
                        ?rest1]]
 [引数   [[ヘッド [[見出し 前]
                    [品詞    <普通名詞>]
                    ?rest2]]
                    ?rest3]]
                    ?rest4]]
                    ?rest5]
:OUT [[ヘッド !x01]
      [LOCT   !x02]
      ?rest5]
)
```

### 2.3 BGRAM (二項規則)

BGRAM は、用言（または <助動詞>「だ/です」）と格（係）助詞の組合せから、最も尤もらしい深層格ラベルを格（係）助詞ノードに付与する規則群である。BGRAM は、手作業で格ラベル

付けを行った格構造データファイルから上記の二項関係の統計を取り、その統計に基づいて作成される。

二項関係の統計例：

```
[[見出し 泊ま]
 [品詞 <本動詞>]
 [活用タイプ 五段ラ]
 [助詞 に]
 [生起回数 29]
 [統計リスト (:LIST (LOCT 25) (TLOC 3) (MANN 1))]]
```

この例は、学習データ中に「～に...泊ま(る)」という表現が29回現われ、格助詞「に」には深層格 LOCT (25回), TLOC (3回), MANN (1回) が与えられていることを示している。

規則例 (BGRAM)：

上記の統計例から生成される規則。統計では、TLOC や MANN より、LOCT の方が頻度が高いので、「～に...泊ま(る)」という表現中の格助詞「に」に対しては LOCT を与えることにする。

```
(defrule
:IN [[ヘッド !x01 [[見出し 泊ま]
                    [品詞 <本動詞>]
                    [活用タイプ 五段ラ]
                    ?rest0]]
     [引数 !x02 [[ヘッド [[見出し に]
                          [品詞 <格助詞>]
                          ?rest1]]
                ?rest2]]
     ?rest3]
:OUT [[ヘッド !x01 ]
      [LOCT !x02 ]
      ?rest3])
```

格助詞は係助詞（「は」「も」など）を介して用言に依存することがあるため、上の規則から派生した次の規則も BGRAM として与えられる。

```
(defrule
:IN [[ヘッド !x01 [[見出し 泊ま]
                    [品詞 <本動詞>]
                    [活用タイプ 五段ラ]
                    ?rest0]]
     [引数 [[ヘッド !x02 [[見出し ?midashi]
                          [品詞 <係助詞>]
                          ?rest1]]
            [引数 !x03 [[ヘッド [[見出し に]
                                  [品詞 <格助詞>]
                                  ?rest2]]
```

```

?rest3]]
    ?rest4]]
?rest5]
:OUT [[ヘッド !x01 ]
    [引数  [[ヘッド !x02 ]
            [LOCT !x03 ]
            ?rest4]]
    ?rest5])

```

注) TGRAM および BGRAM は、依存構造中の素性名 [引数] を深層格ラベルに書き換えている。従って、TGRAM で与えられた深層格ラベルが BGRAM でさらに書き換えられてしまうことはない。

## 2.4 PAT

PAT は <助動詞> 「だ」「です」の直前に来る要素に深層格ラベルを与える規則群である。

<助動詞> 「だ」「です」は、次の構文で用いられる。

- ・ 「A は B です」            A、Bとも名詞類
- ・ 「A は C です」            Cは副詞句
- ・ 「A は D です」            Dは名詞類に助詞類が接続して構成される<後置詞句>
- ・ 「E (～するの)です」      Eは <名詞節> で、ヘッドは <準体助詞> 「の」「ん」。

ここで、PAT が関与するのは B と E であり、PAT 規則により、B には [IDEN]、E には [OBJE] が付与される。A および D には、TGRAM または BGRAM の規則によって深層格ラベルが付与される (A には通常 [OBJE] が与えられる)。副詞類に格は付与しないという原則に基づき、C には深層格ラベルを付与しない。

PAT は出現する可能性のある組合せを網羅するように手作業で作成された。[IDEN] 付与の書き換え規則は、<助動詞> の出現形およびカテゴリと、それに依存するカテゴリとの三項規則である。ここで、<助動詞> 「だ/です」の出現形と<助動詞> 「だ/です」に依存するカテゴリの種類は次のとおりである。

<助動詞> 「だ/です」の出現形 (合計 7 個) :

<助動詞> 「だ」は特殊活用なので、出現形は以下の 6 種類である。

「で」「な」「だろ」「だっ」「だ」「なら」

<助動詞> 「です」の出現形は<助動詞語幹> 「で」だけである。

<助動詞> 「だ/です」に依存するカテゴリの種類 :

B (名詞類) のヘッドのカテゴリは次の 10 個である。

<名詞句><普通名詞><代名詞><サ変名詞><日時><数詞><人名><住所名><接尾辞><準体助詞>

そこで、<助動詞>の出現形と、それに依存するカテゴリの全組み合わせを書き換え規則として用意する (全 70 個)。

```

規則例： (defrule
           :in  [[ヘッド [[見出しで]
                        [品詞 <助動詞>]
                        ?rest1]]
                [引数 [[ヘッド [[品詞 <固有名詞>]
                                ?rest2]]
                        ?rest3]]
                ?rest4]
           :out [[ヘッド [[見出しで]
                          [品詞 <助動詞>]
                          ?rest1]]
                [IDEN [[ヘッド [[品詞 <固有名詞>]
                                ?rest2]]
                       ?rest3]]
                ?rest4]
           )

```

同様に、<準体助詞> (E) の場合は [OBJE] を付与する。

B のカテゴリには <副助詞> が接続して、<名詞句> を構成している可能性がある (<準体助詞> を除く)。その場合、B のヘッドは <副助詞> になるので、以下のような規則群も用意する (全 63 個)。

```

規則例： (defrule
           :in  [[ヘッド [[見出しで]
                        [品詞 <助動詞語幹>]
                        ?rest1]]
                [引数 [[ヘッド [[品詞 <副助詞>]
                                ?rest2]]
                        ?rest3]]
                ?rest4]
           :out [[ヘッド [[見出しで]
                          [品詞 <助動詞語幹>]
                          ?rest1]]
                [IDEN [[ヘッド [[品詞 <副助詞>]
                                ?rest2]]
                       ?rest3]]
                ?rest4]
           )

```

<助動詞> 「だ」「です」に依存している <サ変名詞> は動詞用法でない場合にのみ、深層レベル [IDEN] を付与する。

```

規則例：      (defrule
                :in    [[ヘッド [[見出し で]
                                [品詞 <助動詞語幹>]
                                ?rest1]]
                    [引数 !x1 [[ヘッド [[品詞 <サ変名詞>]
                                        ?rest2]]
                                ?rest3]]
                    ?rest4]
                :env    (no_feature 動詞 !x1) ; 動詞素性がない
                :out    [[ヘッド [[見出し で]
                                [品詞 <助動詞語幹>]
                                ?rest1]]
                    [IDEN [[ヘッド [[品詞 <サ変名詞>]
                                    ?rest2]]
                            ?rest3]]
                    ?rest4]
                )

```

## 2.5 PAT2

<引用助詞>が述語に依存して出現する場合は [CONT] を付与される場合がほとんどである。そこで、<引用助詞> が述語に依存している場合、<引用助詞>の引数について検査することなく、[CONT] を付与する書き換え規則 (PAT2) を手作業で作成した。

```

書き換え規則：      (defrule
                :in    [[引数 [[ヘッド [[品詞 <引用助詞>]
                                        ?rest0]]
                                ?rest1]]
                    ?rest2]
                :out   [[CONT [[ヘッド [[品詞 <引用助詞>]
                                        ?rest0]]
                                ?rest1]]
                    ?rest2]
                )

```

## 2.6 EMPTY

EMPTY は、手作業で省略格付けを行った格構造データファイルから、各用言に関して省略格の統計を取り、生じた用言と省略格の組み合わせすべてについて生成した省略格付与規則である。

省略格の統計例：

```

[[見出し    待]
 [品詞      <本動詞>]
 [活用タイプ 五段タ]
 [統計リスト (:LIST (AGEN 3) (OBJE 2))]]

```

この統計は、学習データ中の動詞「待つ」に関して、省略格 [AGEN NIL] の付与が3回、[OBJE NIL]

の付与が2回あったことを示す。

見つかったすべての省略格ラベルについて EMPTY 規則が作られる。上記の統計からは AGEN と OBJE に対する2つの規則が生成される。

```
(defrule
:IN !x01 [[ヘッド !x02 [[見出し 待]
                        [品詞 <本動詞>]
                        [活用タイプ 五段タ]
                        ?rest0]]
        ?rest1]
:ENV (:no_feature AGEN !x01) ; AGEN が与えられていないという条件
:OUT [[ヘッド !x02]
      [AGEN NIL]
      ?rest1]
)
```

```
(defrule
:IN !x01 [[ヘッド !x02 [[見出し 待]
                        [品詞 <本動詞>]
                        [活用タイプ 五段タ]
                        ?rest0]]
        ?rest1]
:ENV (:no_feature OBJE !x01) ; OBJE が与えられていないという条件
:OUT [[ヘッド !x02]
      [OBJE NIL]
      ?rest1]
)
```

## 2.7 DEL

DEL は、EMPTY の適用により助動詞「だ」「です」に付与された余分な [OBJE NIL] を削除する規則群である。DEL は、用言、助動詞あるいは補助動詞が、助動詞「だ」「です」に依存している場合に、同じレベルに付与された [OBJE NIL] を削除する。

書き換え規則：

<助動詞>「だ」「です」に依存する以下のカテゴリと同じレベルに記述されている [OBJE NIL] を削除する

該当するカテゴリ：

<本動詞> <形容詞> <形容名詞> <サ変名詞><sup>6</sup> <助動詞> <助動詞語幹> <補助動詞>  
<補助動詞語幹> <受身助動詞> <受身助動詞語幹> <使役助動詞> <使役助動詞語幹>

<助動詞> の出現形との組み合わせにより規則数は84個になる。

<sup>6</sup> <サ変名詞> は、動詞用法マーカ ([動詞 +]) が存在する場合にのみ規則が適用される

規則例：

```
(defrule
:in  [[ヘッド [[見出し で]
        [品詞 <助動詞語幹>] ; 「お願いしますでしょうか」で「ます」と同じ
        ?rest1]]
      [OBJE NIL]
      [引数 [[ヘッド [[品詞 <助動詞語幹>]
        ?rest2]]
        ?rest3]]
      ?rest4]
:out [[ヘッド [[見出し で]
        [品詞 <助動詞語幹>]
        ?rest1]]
      [引数 [[ヘッド [[品詞 <助動詞語幹>]
        ?rest2]]
        ?rest3]]
      ?rest4]
)
```

## 2.8 DEL2

DEL2 は、対応する深層格ラベルが係助詞を介して現われている場合に EMPTY によって付与された省略格ラベルを削除する規則である。例えば、EMPTY は動詞「戻る」に「東京には」が依存している場合、[DESTNIL] を付け加える。これは、EMPTY の規則は、深層格ラベル DEST が与えられている格助詞「に」が「は」を介して用言を修飾していることを検知して省略格付けを避けるようにはできていないためである。格助詞「に」に BGRAM によって DEST が与えられている場合、DEL2 の規則は [DESTNIL] を削除する。DEL2 の規則は統計によらずに機械的に生成される。

規則例： (defrule

```
:in  [[OBJE NIL]
      [引数 [[ヘッド [[品詞 <係助詞>]
        ?rest1]]
      [OBJE [[ヘッド [[見出し に]
        [品詞 <格助詞>]
        ?rest2]]
        ?rest3]]
      ?rest4]]
      ?rest5]
:out [[引数 [[ヘッド [[品詞 <係助詞>]
        ?rest1]]
      [OBJE [[ヘッド [[見出し に]
        [品詞 <格助詞>]
        ?rest2]]
        ?rest3]]
      ?rest4]]
      ?rest5]
)
```

## 2.9 ソーティング

ソーティングのプログラム (sort.lisp) は、ヘッドのあるレベルの素性構造を、見出しリスト、品詞リスト、INDEX-LIST、ヘッド、省略格、その他の順に並べ、ヘッドの内部の素性構造を、見出し、INDEX、品詞、活用タイプ、その他の順に並べる（「その他」の要素間のソートは行われない）。このソーティングは、人間にとっての見やすさのために行われる。

## 2.10 REL

関係節の格ラベル付けは、素性構造書き換えシステムではなく専用のプログラムによって行われる。関係節格付与プログラムは Lisp で書かれており、下記に示す例の形式を持った規則 (REL) を読み込み、体言を修飾する用言文節の [引数] ラベルを、最も尤もらしい深層格ラベルに書き換える。関係節の格ラベル付けプログラムが使用する規則 (REL) は、修飾文節の主動詞と被修飾語のバイグラムであり、格構造データファイルからの統計に基づいて作成される。バイグラムは、文節の主たる <本動詞> <サ変名詞> と、文節が連体修飾する語 (ヘッド) との組み合わせに対応する深層格ラベルの頻度を表わす。例えば、次の例は「出る」と「便」が AGEN-REL で依存関係にある頻度を表わしている。

REL 規則 (バイグラム) の例:

```
[[見出し 見]
 [品詞 <本動詞>]
 [活用タイプ 一段]
 [依存先 [[見出し マンハッタン]
          [品詞 <固有名詞>]]]
 [統計リスト (:LIST [OBJE-REL 3])]
]]
```

この例は、学習データのうち、動詞「見る」を主動詞とするような文節が固有名詞「マンハッタン」にかかるような例では、「マンハッタン」の深層格が OBJE-REL であるようなものが3件あったことを示している。(実際に実験に用いたデータからは、統計リストの要素が1個より多いような規則は得られなかった。) この場合、「見る」を主たる動詞とする文節が「マンハッタン」を修飾する場合に、OBJE-REL が最もしばしば現われているので、こうした場合のデフォルト深層格は OBJE-REL となる。

<形容詞>、<形容名詞> が修飾語である場合、関係詞節デフォルト格ラベル付けプログラムは OBJE-REL を関係詞節デフォルト格ラベルとして与える。

## 3 助詞類の出現パターンによる処理方法

半自動格付与プログラムは、<格助詞>、<係助詞>、<引用助詞> を格マーカと認定し、書き換え規則を作成する。格付与の際に問題となった助詞の諸現象への対処を以下に述べる。

### 3.1 助詞の連続

<格助詞> には、強調の意味で <係助詞> が接続することが多い。<格助詞> が <係助詞> を介して用言に依存する場合、<格助詞> と用言に直接依存する場合の深層格ラベルの付与と同等に扱って、<格助詞> に深層格ラベルを付与する。この操作を行う規則は <格助詞> と用言から最も尤もらしい深層格を決定する規則から派生的に生成され、BGRAM 規則として適用される (2.3 参照)。



## 3.2 助詞の交替

コントロール現象や強調表現による助詞の交替については、現在は、出現した助詞と格ラベルの組合せをそのまま規則化している。

### 3.2.1 コントロール現象による助詞の交替

受身、使役のコントロール現象により交替した助詞には深層格が付与されており、それらは能動文と同じ扱いで統計データとして抽出され、書き換え規則になっている。その結果、格マーカが一致すると、誤った格が付与される場合が生じる。以下、「行なう」と共起した格マーカ「を」と「行なわれる」と共起した格マーカ「が」についてどのような現象が起きるかを例示する。

TGRAM 用統計:

```
[[見出し 行な] ;;<-- 能動態
 [品詞 <本動詞>]
 [活用タイプ 五段ワ]
 [助詞 を]
 [目的語見出し 予約]
 [目的語品詞 <サ変名詞>]
 [生起回数 1]
 [統計リスト (:LIST (OBJE 1))]]
```

```
[[見出し 行な] ;;<-- 受動態
 [品詞 <本動詞>]
 [活用タイプ 五段ワ]
 [助詞 が]
 [目的語見出し 変更#手続き]
 [目的語品詞 <名詞句>]
 [生起回数 1]
 [統計リスト (:LIST (OBJE 1))]]
```

BGRAM 用統計:

```
[[見出し 行な] ;;<-- 能動態
 [品詞 <本動詞>]
 [活用タイプ 五段ワ]
 [助詞 を]
 [生起回数 1]
 [統計リスト (:LIST (OBJE 1))]]
```

```
[[見出し 行な] ;;<-- 受動態
 [品詞 <本動詞>]
 [活用タイプ 五段ワ]
 [助詞 が]
 [生起回数 1]
 [統計リスト (:LIST (OBJE 1))]]
```

EMPTY 用統計：（受動態と能動態の格パターンが統合される）

```
[[見出し 行な]
 [品詞 <本動詞>]
 [活用タイプ 五段ワ]
 [統計リスト (:LIST (AGEN 5) (OBJE 1))]]
```

「は」「も」に付与された格ラベル

上の統計から生成した規則群により「彼が予約を行なった」に格付与したとすると、次のような出力結果が予想できる。

```
「彼が 予約を 行なった」
[AGEN NIL] [OBJE] [OBJE]
(省略格)
```

ここで、[AGEN NIL] は不要であり、「彼が」に対する [OBJE] 付与は間違い（[AGEN] が正しい）である。

また、動詞「含む」については、能動態と受動態とでは、手作業で格付与をしていた時期から異なる格パターンを付与していた（ASURA の格付与基準に基づいていたため）。

```
「その料金は 税金を 含みます」
[AGEN] [OBJE]
```

```
「その料金には 税金が 含まれます」
[LOCT] [OBJE]
```

その結果、EMPTY 用の統計は次のようになる。

```
[[見出し 含]
 [品詞 <本動詞>]
 [活用タイプ 五段マ]
 [統計リスト (:LIST (LOCT 6) (OBJE 3) (AGEN 3))]]
受動態の場合 共通 能動態の場合
```

そして、「含む」が述語である受身の文に格付与を行なった場合、次のような結果が出力される。

```
「宿泊料金に含まれておりません」
[AGEN NIL] [OBJE NIL] [LOCT]
省略格 省略格
^^^^^^ <- 不要
```

このような現象は、規則化の過程で、能動態と受動態の場合の処理を分離することによって解決される。

### 3.2.2 強調による助詞の交替

強調により助詞が交替した場合、例えば BGRAM に次のような書き換え規則が登録される。

BGRAM 用統計:

[[見出し ほし]

[品詞 <形容詞>]

[助詞 が]

[生起回数 5]

[統計リスト (:LIST (OBJE 4)(引数 1))]]

[[見出し ほし]

[品詞 <形容詞>]

[助詞 も]

[生起回数 1]

[統計リスト (:LIST (OBJE 1))]]

ここで、述語「ほしい」に依存する<係助詞>「も」は、<格助詞>「が」(OBJE)の交替である。上の規則により、「僕もお菓子がほしい」という文に格付与したとすると、次のような出力結果が予想できる。

「僕も お菓子が ほしい」

[EXPR NIL] [OBJE] [OBJE]

省略格

ここで「僕も」には [EXPR] が付与されるべきであり、[EXPR NIL] は不要である。係助詞は「が」「を」と交替する。多くの場合、この問題は TGRAM を用いることにより解決できる。

### 3.3 助詞の省略

格付与プログラムでは、格マーカとなる助詞が省略された名詞類を格付与対象語句として扱わない。

例：わたし、鈴木弘子といひます。 ; 「わたし」に接続する助詞の省略

三項規則による格付与は可能であるが、実装していない。

### 3.4 <副助詞>の扱い

名詞類に<副助詞>が接続する場合、構文解析の段階では2通りの処理を行なっている。

1. <副助詞>が接続して<名詞句>を構成する。

(<名詞句> <-> (<名詞句> <副助詞>)) ; <助動詞>「だ」「です」が接続する場合

2. <副助詞>が接続して<後置詞句>を構成する。

(<後置詞句> <-> (<名詞句> <副助詞>)) ; その他の述語を修飾する場合

手作業では、2. の場合に対して格ラベルを付与しているが、半自動格付与では、<副助詞> を格マーカと認めないので、書き換え規則が作成されない。

例：柿なんて食べられません。

書き換え規則による格付与後：柿なんて [AGEN NIL] [OBJE NIL] 食べられません。

↓

省略格

手作業による修正後：柿なんて [AGEN NIL] 食べられません。

[OBJE] 省略格

## 4 実験結果

### 4.1 評価の概要

依存構造に格ラベルを上記の方法で自動付与し、人手による正解データと比較することにより、格ラベル付与の評価を行なった。TGRAM、BGRAM 等の統計データを抽出するための学習データとしては、人手によって格ラベル付与を行ったSLDBの一部(375ファイル)を用いた。格ラベル自動付与対象の依存構造は、統計学習用の375ファイルとは重ならないSLDBの依存構造データ(119ファイル<sup>7</sup>、3364文)である(付録2参照)。

### 4.2 評価の方法

前述の二つの格ラベル付き依存構造をFS1(書き換え結果データ)、FS2(正解データ)とし、FS2の深層格ラベルの数とそのラベルに対応するFS1のラベルが一致するもの(下記参照)の数を数える。ここで再現率は次の式で与えられる。

再現率 =

$$\frac{\text{FS1の深層格ラベルがFS2の深層格ラベルと一致した数}}{\text{FS2の深層格ラベルの数}} * 100$$

次に、FS1の深層格ラベルの数とそのラベルに対応するFS2のラベルが一致するもの(下記参照)の数を数える。ここで適合率は次の式で与えられる。

適合率 =

$$\frac{\text{FS2の深層格ラベルがFS1の深層格ラベルと一致した数}}{\text{FS1の深層格ラベルの数}} * 100$$

この時、再現率、適合率ともに100%であるデータを文における正解とする。

FS1、FS2の2つのノードの対応は次の条件によって判定した。

- ・省略格ラベルの場合  
同じ省略格ラベルを持ち、値がNILである。
- ・その他の格ラベルの場合  
同じ省略格ラベルを持ち、サブレベルの次の素性の値すべてが一致する。
  - ・品詞リスト
  - ・見出しリスト
  - ・INDEX-LIST

### 4.3 評価結果

前述の評価方法を用い、オープンデータ119ファイルの深層格自動付与結果に対して評価を行なった。以下、総合評価につづいて、格助詞、係助詞に関する結果、省略格、関係節格についての結果、さらに述語要素の品詞別による結果を示す。

#### ・総合評価

文の数	: 3364
正解の文の数	: 1347

<sup>7</sup> 375ファイルは1996年3月までに格ラベル付与が終了していた格構造ファイルのすべてであり、119ファイルは1996年9月までに格ラベル付与が終了していない依存構造ファイルのすべてである。

文正解率	: 40.04 %
正解データのアークラベルの総数	: 9857
正解データのアークラベルが一致した数	: 5835
再現率	: 59.2 %
結果データのアークラベルの総数	: 8547
結果データのアークラベルが一致した数	: 6177
適合率	: 72.27 %

・ 正解では深層格ラベルがついているのに、最終結果ファイルでは「引数」になっているものの数と割合（省略格はカウントの対象外）

正解データのアークラベルの総数	: 5015
正解データのアークラベルが一致した数	: 3741
再現率	: 74.6 %
結果データのアークラベルの総数	: 4333
結果データのアークラベルが一致した数	: 3741
適合率	: 86.34 %
結果データのアークラベルが引数の数	: 902
引数のままであった割合	: 17.99 %

・ 格助詞

正解データのアークラベルの総数	: 2535
正解データのアークラベルが一致した数	: 1754
再現率	: 69.19 %
結果データのアークラベルの総数	: 2100
結果データのアークラベルが一致した数	: 1754
適合率	: 83.52 %

・ 格助詞（TGRAM のみの適用による結果）

正解データのアークラベルの総数	: 2535
正解データのアークラベルが一致した数	: 691
再現率	: 27.26 %
結果データのアークラベルの総数	: 712
結果データのアークラベルが一致した数	: 691
適合率	: 97.05 %

格助詞の深層格付与において、TGRAM のみの適用結果では適合率が高く、パターンがマッチすればほぼ正解である。しかし再現率は低く、データがスパースであることを示している。一方、BGRAM を経たデータでは、再現率が上がっているが、適合率は低下しており、格要素（体言）の情報がないと、正確な深層格付与ができないことを示している。

・ 係助詞

正解データのアークラベルの総数	: 580
正解データのアークラベルが一致した数	: 414
再現率	: 71.38 %

結果データのアークラベルの総数	: 573
結果データのアークラベルが一致した数	: 414
適合率	: 72.25 %

・省略格

正解データのアークラベルの総数	: 4842
正解データのアークラベルが一致した数	: 2094
再現率	: 43.25 %
結果データのアークラベルの総数	: 4214
結果データのアークラベルが一致した数	: 2436
適合率	: 57.81 %

・関係節

正解データのアークラベルの総数	: 158
正解データのアークラベルが一致した数	: 78
再現率	: 49.37 %
結果データのアークラベルの総数	: 121
結果データのアークラベルが一致した数	: 78
適合率	: 64.46 %

→関係節補足実験（次頁）参照

・本動詞、サ変名詞（動詞用法）<sup>8</sup>

正解データのアークラベルの総数	: 6582
正解データのアークラベルが一致した数	: 4269
再現率	: 64.86 %
結果データのアークラベルの総数	: 5285
結果データのアークラベルが一致した数	: 4279
適合率	: 80.96 %

・形容詞、形容名詞

正解データのアークラベルの総数	: 565
正解データのアークラベルが一致した数	: 434
再現率	: 76.81 %
結果データのアークラベルの総数	: 605
結果データのアークラベルが一致した数	: 434
適合率	: 71.74 %

<sup>8</sup> 関係節の統計は含まない。また、用言に直接格要素が依存していない場合（係助詞を介している場合）もカウントしない。次の形容詞の統計も同様である。

## ・だ/です

正解データのアークラベルの総数	: 2378
正解データのアークラベルが一致した数	: 2287
再現率	: 96.17 %
結果データのアークラベルの総数	: 2411
結果データのアークラベルが一致した数	: 2287
適合率	: 94.86 %

## ・関係節補足実験

関係節の用言と係り先のペアがトレーニングデータ中にない場合に、見出しを考慮しない統計を用いてラベル付けを行った。

正解データのアークラベルの総数	: 158
正解データのアークラベルが一致した数	: 92
再現率	: 58.23 %
結果データのアークラベルの総数	: 194
結果データのアークラベルが一致した数	: 92
適合率	: 47.42 %

予想されるとおり、元の結果ファイルよりも多くの関係節格ラベルが付加され、再現率が上昇しているが、適合率は低下している。これは、係り先の情報がないと正確な深層格付与ができないことを示している。

## 5 問題点と改良案

現段階での書き換えシステムおよび規則に関する問題点と改良案を以下に挙げる。

### 5.1. 関係節の場合の格の二重付与

関係節の深層格付与プログラムは、既に与えられた深層格ラベルをチェックしていないため、それまでに TGRAM, BGRAM, あるいは EMPTY により与えられているものと同じ深層格を二重に与えてしまうことがある。この問題は、関係節の深層格付与プログラムを改良することにより解決することができる。特に、統計データに深層格の第2候補以下が存在する場合には、既に与えられていない深層格を選んで付与するようにすることもできる。

例：ワシントンホテル五〇七号室に滞在しております田中と申します

```
[引数 [[見出しリスト (:LIST ま  
す)]  
[品詞リスト (:LIST <助動詞語幹>  
<語尾>)]  
[INDEX-LIST (:LIST 15  
16)]  
[ヘッド [[見出し ま  
[INDEX 15]  
[品詞 <助動詞語幹>]]]  
[引数 [[見出しリスト (:LIST 申  
し)]  
[品詞リスト (:LIST <本動詞>  
<語尾>)]
```



それが <助動詞> に依存している場合、<連用修飾> のヘッドは <形容名詞> なので、<助動詞> に <形容名詞> が依存している形になり、DEL が適用されて必要な [OBJE NIL] が削除される。実際には、このような現象は稀である。

例：確かにそうでございますが…

```
[引数 [[見出しリスト (:LIST ございま
          す)]
      [品詞リスト (:LIST <補助動詞語幹>
                    <語尾>)]
      [INDEX-LIST (:LIST 7
                     8)]
      [ヘッド [[見出し ございま]
                [INDEX 7]
                [品詞 <補助動詞語幹>]]]
      [引数 [[見出しリスト (:LIST で)
                [品詞リスト (:LIST <助動詞>)]
                [INDEX-LIST (:LIST 6)]
                [ヘッド [[見出し で]
                          [INDEX 6]
                          [品詞 <助動詞>]]] ; <-- <助動詞>は A)の<形容名詞>を参照している
                [OBJE NIL] ; 必須格だが、削除されてしまう
      [引数 [[見出しリスト (:LIST そう)]
                [品詞リスト (:LIST <副詞>)]
                [INDEX-LIST (:LIST 5)]
                [ヘッド [[見出し そう]
                          [INDEX 5]
                          [品詞 <副詞>]]]]]
      [引数 [[見出しリスト (:LIST 確か
                            に)]
                [品詞リスト (:LIST <形容名詞>
                              <助動詞>)]
                [INDEX-LIST (:LIST 3
                              4)]
                [ヘッド [[見出し 確か]
                          [INDEX 3]
                          [品詞 <形容名詞>]]]]]]]] ; <-- A)
```

### 5.3 EMPTY に登録されている省略格パターンの整備

書き換え規則作成では、まず、述語に対する格パターンを抽出するが、EMPTY に変換する際には、述語が同じであれば、多義語の場合でも、複数の格パターンを1個に統合している。

例えば、「なる」の場合、以下の5つの格パターンが存在する。

語義 1: [OBJE]	;; [朝食は]どうなりますか
語義 2: [OBJE][IDEN]	;; [料金は][一泊二万円に]なります
語義 3: [OBJE][RESL]	;; [京都に行くことに]なりました
語義 4: [AGEN]	;; [わたしは]横になる
語義 5: [EXPR][CONT][RESL]	;; [わたしは][~かどうか][気に]なる

EMPTY は、これを統合するので、以下のようなデータが作成される。

```

[[見出し   な]
 [品詞     <本動詞>]
 [活用タイプ 五段ラ]
 [統計リスト (:LIST (OBJE 392) (RESL 2) (EXPR 1) (AGEN 1) (IDEN 1))]]

```

その結果、EMPTY の適用後、次のような結果が出力される。

例：一泊九十ドルとなっております

```

[引数 [[見出しリスト (:LIST ま
                               す)]]
 [品詞リスト (:LIST <助動詞語幹>
                   <語尾>)]
 [INDEX-LIST (:LIST          17
                18)]
 [ヘッド [[見出し ま]
           [INDEX          17]
           [品詞 <助動詞語幹>]]]
 [引数 [[見出しリスト (:LIST てお
                               り)]]
 [品詞リスト (:LIST <助動詞語幹>
                   <語尾>)]
 [INDEX-LIST (:LIST          15
                16)]
 [ヘッド [[見出し てお]
           [INDEX          15]
           [品詞 <助動詞語幹>]]]
 [引数 [[見出しリスト (:LIST な
                               っ)]]
 [品詞リスト (:LIST <本動詞>
                   <語尾>)]
 [INDEX-LIST (:LIST          13
                14)]
 [ヘッド [[見出し な]
           [INDEX          13]
           [品詞 <本動詞>]
           [活用タイプ 五段ラ]]]
 [IDEN [[見出しリスト (:LIST と)]]
 [品詞リスト (:LIST <格助詞>)]
 [INDEX-LIST (:LIST          12)]
 [ヘッド [[見出し と]
           [INDEX          12]
           [品詞 <格助詞>]]]
 [引数 [[見出しリスト (:LIST 一
                               泊
                               九#十
                               ドル)]]
 [品詞リスト (:LIST <数詞>
                   <接尾辞>
                   <数詞>
                   <接尾辞>)]
 [INDEX-LIST (:LIST          8
                9
                10
                11)]
 [ヘッド [[見出し ドル]
           [INDEX          11]
           [品詞 <接尾辞>]]]]]]]
 [AGEN NIL] ;; <-- 不要な省略格
 [EXPR NIL] ;; <-- 不要な省略格
 [RESL NIL] ;; <-- 不要な省略格
 [OBJE NIL]]]]]]] ;; <-- 正しい省略格

```

また、以前に誤って付与した格ラベル、あるいは、使役表現のために付与した格ラベルが、統計データとして抽出され、書き換え規則として適用されている例がある。

「待つ」 … [RECP]

EMPTY:

[[見出し 待]

[品詞 <本動詞>]

[活用タイプ 五段タ]

[統計リスト (:LIST (AGEN 438) (OBJE 373) (RECP 1))]]

^^^^^^ <-- 使役のため

これらも「なる」と同様に、不要な省略格が付与される原因になる。

これを防ぐため、次の方法が考えられる。

- 1) 格パターン抽出→格パターン統合 のあとに、EMPTY を人手で整備する

[[見出し な]

[品詞 <本動詞>]

[活用タイプ 五段ラ]

[統計リスト (:LIST (OBJE 392) (RESL 2) (EXPR 1) (AGEN 1) (IDEN 1))]]

生起回数2回の [RESL]、生起回数1回の [EXPR]、[AGEN]、[IDEN] を削除対象とする。

- 2) 格パターン抽出後に複数の格パターンを1つに統合せず、最も出現頻度が多い格パターンを当該の述語の省略格パターンとする。

語義1: [OBJE] ; ; 朝食はどうなりますか

語義2: [OBJE][IDEN] ; ; 料金は一泊二万円になります

語義3: [OBJE][RESL] ; ; 京都に行くことになりました

語義4: [AGEN] ; ; 横になる

語義5: [EXPR][RESL][CONT] ; ; ~かどうかが気になる

「なる」については、語義1 が最も出現頻度が多い (と思われる) ので、それを省略格パターンとする。例えば、次のようになる。

[[見出し な]

[品詞 <本動詞>]

[活用タイプ 五段ラ]

[統計リスト (:LIST (OBJE 392))]]

## 5.4 データサイズとスパース性

実験結果は TGRAM の有効性を示しているが、TGRAM のデータは膨大なものになる上 (実験用の TGRAM で 5 MB 強)、オープンデータではスパース性の問題が起こり、規則があたりにくい。REL も 2 つの自立語を二項関係の要素としているために、スパース性の問題がある。

以下に、格付与におけるスパース性の問題に関して有効と思われる手法を挙げる。

### 5.4.1. 見出しの一部を用いる方法

単語の見出しの後部が意味的な情報を持つと考え、見出し語の後方部分一致によるパターンマッチングを行う。例えば、「京都駅から三十分で着きました。」という文に深層格付与を行う場合、学習データに「大阪駅 から 着く」というパターンはあるが、「京都駅 から 着く」というパター

ンはないとしよう。この場合、次頁の例のような見出しの後方部分一致による書き換え規則と、それを実行する機構を作成すれば、適切な深層格付与を行うことができる。こうした規則の生成は例えば「京都駅」を「京都」と「駅」の2つの形態素に分解するプログラムを用いることにより、ある程度自動化できると考えられる。

```

例：
(defrule
:IN [[ヘッド !x01 [[見出し 着]
      [品詞 <本動詞>]
      [活用タイプ 五段カ]
      ?rest0]]
  [引数 !x02 [[ヘッド [[見出し から]
      [品詞 <格助詞>]
      ?rest1]]
  [引数 [[ヘッド [[見出し ****駅] ;; <-- 後方一致で、末尾が「駅」
      [品詞 <固有名詞>] ;; という文字列に適用
      ?rest2]] ;; **** は任意の文字列を表す。
      ?rest3]]
      ?rest4]]
  ?rest5]
:OUT [[ヘッド !x01 ]
      [DEPT !x02 ] ;; <-- 格ラベルは一意に決まる
      ?rest5] )

```

#### 5.4.2. 意味情報を用いる方法

TGRAM 中の名詞について、見出しではなく意味素性あるいはシソーラスの意味分類を用いることにより、データ量とスパース度を減少させることができると考えられる。この方法のためには、意味情報を持った名詞辞書が必要である。さらに名詞に意味素性を与えた格構造データファイルによって TGRAM を生成することが望ましい。

規則例：「～にある」というパターンで「～」の部分が意味素性 [LOC +] を持っていれば、「に」に深層格 [LOCT] を与える三項関係規則

```

(defrule
:IN [[ヘッド !x01 [[見出し あ]
      [品詞 <本動詞>]
      [活用タイプ 五段ラ]
      ?rest0]]
  [引数 !x02 [[ヘッド [[見出し に]
      [品詞 <格助詞>]
      ?rest1]]
  [引数 [[ヘッド [[LOC +] ;; <- 意味素性
      [品詞 <普通名詞>]
      ?rest2]]
      ?rest3]]
      ?rest4]]
  ?rest5]

```

```
:OUT [[ヘッド !x01 ]  
      [LOCT !x02 ] ;; <- 深層格付与  
      ?rest5 ] )
```

上記の例に見るように、この方法は深層格 TLOC と LOCT の多義性解消に役立つ。ただし、2.2 の例「前にある」のように、名詞部に多義性があるような場合は、この方法でも深層格の多義性解消をすることができない。

#### 参考文献

- [1] 浦谷則好, 竹澤寿幸, 松尾秀彦, 森田千帆, "音声言語データベースの構成" TR-IT-0056, エイ・ティ・アール音声翻訳通信研究所, 1994.5
- [2] 荒川直哉 "ATR 音声言語データベース日本語依存構造・格構造データファイルの作成" TR-IT-0220, エイ・ティ・アール音声翻訳通信研究所, 1994.5

#### 謝辞

デフォルトグラベル自動付与実験関連のプログラミングと作業を行っていただいた谷田泰郎氏および、定性的な評価を行っていただいた日本アイアールの坂口明子氏にこの場を借りて感謝いたします。

## 付録

### 1 深層格のリスト

格構造データファイルで用いられており、今回の実験で付与対象となった深層格ラベルは次の通り。

[深層格]	[定義]	[例]
AGEN	有生・無生の行為主体	私が 会議に申し込む 計算機が 文を解析する
EXPR	精神的事象の体験者 被害を受ける者	母が 悲しんだ 弘が 金をなくした
OBJE	変化・移動・行為の対象 状態動詞の主体	申込用紙を 送る 会議が ある
SOUR	変化の始状態	父が 癌から 快復した 信号が 赤から 変わった
RESL	変化の結果	委員長に なる 信号が 赤へ 変わった
GOAL	行為が向かうもの	木に ぶつかる
SEPR	分離の作用の相手	危険から 身を避ける 水から 酸素を分離する
MUTL	相互動詞の相手	論敵と 対立する
RECP	所有移動における受け手	登録用紙を 事務局に 送る
ORIG	所有移動における与え手	登録用紙を 事務局から 送る
CONT	思考・発言などの内容	妻を同伴しようと思ふ
ROUT	移動における通過経路	公園を 通る 橋を 渡る
MANN	様態	厳しい態度で 臨む
INST	方法・手段	ワープロで 書く
MATR	材料・構成要素	葡萄から ワインを作る 3章から なる論文
CAUS	原因・理由	癌で 死ぬ
PURP	目的	買い物に 行く
COND	条件	無料で 招待する
LOCT	場所	登録用紙に 記入する 彼の意見には 偏見がある
DEPT	起点	京都駅から 出発する この風習は 仏教から きた
DEST	終点	そちらの住所に 送る
TLOC	時点	会議の際に 発表する
TDEP	時間的起点	9時から 始まる
TDES	時間的終点	月末までに 払う
ROLE	役割	講演者として 参加する
ACCM	随伴者	友達と 映画に行く
RANG	範囲規定	内容について 質問する
COMP	比較	ビールより ワインがいい
IDEN	同一	彼の作品が 金賞に なった
TOPIC	トピック	私は 背が 高い

・提題格 TOPIC は深層格ではなく「動詞辞書 (Ver.2)」にも記載されていないが、構文上の役割として導入されている。

## 2 ファイルリスト

・学習データ用ファイル (375個)

TAC22011	TAS22021	TBS33001	TCS12023	TCS33015	THS32002	TOS33004
TAC22012	TAS22022	TBS33002	TCS12024	TCS33025	THS32003	TOS33005
TAC22013	TAS22031	TBS33003	TCS13001	TCS33026	THS32004	TOS33006
TAC22014	TAS22032	TBS33004	TCS13002	TCS33027	THS33001	TRS12001
TAC23011	TAS23001	TCC22011	TCS13003	TCS33028	THS33002	TRS12002
TAC23012	TAS23002	TCC22012	TCS13004	TCS33029	TIS12001	TRS13001
TAC23013	TAS23001	TCC22013	TCS13005	TCS33030	TIS12002	TRS13002
TAC23014	TAS23002	TCC22014	TCS13006	TCS33031	TIS13001	TRS22001
TAC23021	TAS23003	TCC22031	TCS13017	TCS33032	TIS13002	TRS22002
TAC23022	TAS23004	TCC22032	TCS13018	TCS33033	TIS22001	TRS23001
TAC23023	TAS23005	TCC22033	TCS13019	TCS33034	TIS22002	TRS23002
TAC23024	TAS23006	TCC22034	TCS13020	TDS12003	TIS23001	TRS32001
TAC23031	TAS23007	TCC22041	TCS13021	TDS12004	TIS23002	TRS32002
TAC23032	TAS23008	TCC22042	TCS13022	TDS12005	TIS32001	TRS33001
TAC23033	TAS23009	TCC22043	TCS13023	TDS12006	TIS32002	TRS33002
TAC23034	TAS23010	TCC22044	TCS13024	TDS12007	TIS33001	TSS12001
TAS12001	TAS32011	TCC22071	TCS22021	TDS12008	TIS33002	TSS12002
TAS12002	TAS32012	TCC22072	TCS22022	TDS12009	TKS12001	TSS13001
TAS12005	TAS32013	TCC22073	TCS22051	TDS12010	TKS12002	TSS13002
TAS12006	TAS32014	TCC22074	TCS22052	TDS13005	TKS13001	TSS22001
TAS12007	TAS32015	TCC22091	TCS22061	TDS13006	TKS13002	TSS22002
TAS12008	TAS32016	TCC22092	TCS22062	TDS13007	TKS22001	TSS23001
TAS12009	TAS33001	TCC22093	TCS22081	TDS13008	TKS22002	TSS23002
TAS12010	TAS33002	TCC22094	TCS22082	TDS13009	TKS23001	TSS23001
TAS12011	TAS33009	TCC23011	TCS32001	TDS13010	TKS23002	TSS23002
TAS12012	TAS33010	TCC23012	TCS32002	TDS32005	TKS32001	TSS33001
TAS12013	TAS33011	TCC23013	TCS32003	TDS32006	TKS32002	TSS33002
TAS12014	TAS33012	TCC23014	TCS32004	TDS32007	TKS33001	
TAS12015	TAS33013	TCC23021	TCS32005	TDS32008	TKS33002	
TAS12016	TAS33014	TCC23022	TCS32006	TDS33007	TOS12001	
TAS12017	TAS33015	TCC23023	TCS32007	TDS33008	TOS12002	
TAS12018	TAS33016	TCC23024	TCS32008	TDS33009	TOS12003	
TAS12019	TAS33017	TCC23031	TCS32009	TDS33010	TOS12004	
TAS12020	TAS33018	TCC23032	TCS32010	TGS12001	TOS13001	
TAS12021	TAS33019	TCC23033	TCS32011	TGS12002	TOS13002	
TAS12022	TAS33020	TCC23034	TCS32012	TGS13001	TOS13003	
TAS13001	TAS33021	TCC23041	TCS32013	TGS13002	TOS13004	
TAS13002	TAS33022	TCC23042	TCS32014	TGS22001	TOS13005	
TAS13005	TBS12001	TCC23043	TCS32015	TGS22002	TOS13006	
TAS13006	TBS12002	TCC23044	TCS32016	TGS23001	TOS22001	
TAS13007	TBS12003	TCC23051	TCS32017	TGS23002	TOS22002	
TAS13008	TBS12004	TCC23052	TCS32018	TGS32001	TOS22003	
TAS13009	TBS13001	TCC23053	TCS32019	TGS32002	TOS22004	
TAS13010	TBS13002	TCC23054	TCS32020	TGS33001	TOS23001	
TAS13011	TBS13003	TCC23061	TCS32021	TGS33002	TOS23002	
TAS13012	TBS13004	TCC23062	TCS32022	THS12001	TOS23003	
TAS13013	TBS22001	TCC23063	TCS32023	THS12002	TOS23004	
TAS13014	TBS22002	TCC23064	TCS32024	THS12003	TOS23005	
TAS13015	TBS22003	TCS12001	TCS33001	THS12004	TOS23006	
TAS13016	TBS22004	TCS12002	TCS33002	THS13001	TOS32001	
TAS13017	TBS23001	TCS12003	TCS33003	THS13002	TOS32002	
TAS13018	TBS23002	TCS12004	TCS33004	THS22001	TOS32003	
TAS13019	TBS23003	TCS12005	TCS33009	THS22002	TOS32004	
TAS13020	TBS23004	TCS12006	TCS33010	THS22003	TOS32005	
TAS13021	TBS32001	TCS12019	TCS33011	THS22004	TOS32006	
TAS13022	TBS32002	TCS12020	TCS33012	THS23001	TOS33001	
TAS22001	TBS32003	TCS12021	TCS33013	THS23002	TOS33002	
TAS22002	TBS32004	TCS12022	TCS33014	THS32001	TOS33003	

・格付与実験対象ファイル (119 個)

TAS12023	TCC23091	TCS12033	TCS23131	TCS33037	TCS33057
TAS12024	TCC23092	TCS12034	TCS23132	TCS33038	TCS33058
TAS12025	TCC23093	TCS12035	TCS23141	TCS33039	TCS33059
TAS12026	TCC23094	TCS12036	TCS23142	TCS33040	TCS33060
TAS13023	TCC23101	TCS12037	TCS23151	TCS33041	TCS33061
TAS13024	TCC23102	TCS12038	TCS23152	TCS33042	TCS33062
TAS13025	TCC23103	TCS13029	TCS32029	TCS33043	TCS33063
TAS13026	TCC23104	TCS13030	TCS32030	TCS33044	TCS33064
TCC22111	TCC23111	TCS13031	TCS32031	TCS33045	TIS33010
TCC22112	TCC23112	TCS13032	TCS32032	TCS33046	TIS33011
TCC22113	TCC23113	TCS13033	TCS32033	TCS33047	TIS33012
TCC22114	TCC23114	TCS13034	TCS32034	TCS33048	TOS32010
TCC23071	TCC23121	TCS13035	TCS32035	TCS33049	TOS33010
TCC23072	TCC23122	TCS13036	TCS32038	TCS33050	TOS33011
TCC23073	TCC23123	TCS13037	TCS32039	TCS33051	TOS33012
TCC23074	TCC23124	TCS13038	TCS32040	TCS33052	TOS33013
TCC23081	TCS12029	TCS22101	TCS32041	TCS33053	TOS33014
TCC23082	TCS12030	TCS22102	TCS32042	TCS33054	TOS33015
TCC23083	TCS12031	TCS22121	TCS33035	TCS33055	TOS33016
TCC23084	TCS12032	TCS22122	TCS33036	TCS33056	