

日本語文間の意味関係解析システム InSeRA の開発研究

川端 崇央 * 原田 実 **

青山学院大学 理工学研究科 経営工学専攻 *

青山学院大学 理工学部 情報テクノロジー学科 **

文間の理由, 条件, 時間的近接などの意味的な関係を表す 21 個の文間深層格を定義し, 実際に日本語文章を処理し, 隣接 2 文間毎に文間深層格を記述したリストを自動生成するシステム InSeRA を開発した. ここでは, 隣接 2 文間に, 「だから～」のような接続表現など表層的な手がかりや助詞の活用形, EDR 電子化辞書による, 文中の中心用言の概念 ID などを用いたヒューリスティックな規則を適用することによって, 2 文間の文間深層格を推定する手法を提案する.

Development research on system InSeRA which analyzes the semantic relations between Japanese sentences

Takao Kawabata * Minoru Harada **

Graduate School of Industrial and System Engineering, Department of Science and Engineering,
Aoyama Gakuin University *

Department of Integrated Information Technology, Faculty of Science and Engineering,
Aoyama Gakuin University **

In this research, in order to clarify the semantic relations between sentences such as “reason” “condition” and “time adjacent”, we defined 21 kinds of “Deep case between sentences”. And also we developed the system InSeRA, which automatically analyzes these deep cases between Japanese sentences. For this purpose, InSeRA applies to the semantic frames of sentence S_i and the semantic frames of sentence S_{i+1} the heuristic rules based on the conjunction, the word meaning devised from the EDR electronic dictionary, and the inflection of the auxiliary verb.

1. はじめに

1.1. 研究の背景

我々はこれまで, 日本語文章を解析し, 文中の語の意味と語間の格関係を意味フレーム(格フレーム)形式として自動生成する意味解析シ

ステム SAGE[5]を開発してきた. しかし, 本研究室で進められているオブジェクト指向設計の動的分析(イベント時系列の解析)の自動化や判例文検索における事故状況の類似度解析(用言間の時間的, 論理的関係の解析)などの応用を考えると, 文をまたがった語の意味的關係の解

析が必要となる。これは一般に文脈解析と呼ばれる解析の一部である。

一般的に文脈解析とは、文と文との間の関係をとらえることにより、文によって表されている概念や事象のつながりを明らかにし、最終的には結束性、つまり文と文の間をまたがって各語の指す概念や事象の意味的な関係をすべて明らかにすることを目的とする。ここには文章の背後にある概念や事象の間の関係も含まれている。ここで、文間の関係は、長尾氏、角田氏らがまとめた「自然言語処理」の第7章「文脈解析」[1]において述べられているものであり、結束性はボウグランド(R.de.Beaugrande)、ドレスラー(W.U.Dressler)の両氏が「テキスト言語学入門」[2]において、提唱したものである。

個々の概念や事象の間の関係をコンピュータによって厳密に求めるためには、いろいろな手がかりを使い試行錯誤的な推論やパターン認識を行っている人間の思考過程をもっとよく検討し、その上でそれを規則や知識にし、コンピュータに実装しなければならない。しかし、そのようなものすべてが明らかにされ実装されなくても、表層的な手がかりを用いて文間の意味的な関係のある程度明らかにすることはできる。逆に文章の意味的な構造を表現するときには、そのような明示的な表現が本質的である。また、文と文のつながりをとらえることは、論説文やシステムの要求仕様文の場合などでは特に、論理の流れをとらえる際の重要な指標となる。

文間の意味的な関係解析の研究としては、田中、林[7]らが、ソフトウェアのマニュアルを例にとり、記述される世界の特徴を反映した文のタイプ分類及びその並びのパターン(接続パターン)に基づいて隣接する2文間の関係を解析する手法を提案している。我々が「。」で区切られた1文1文を1つのノードとして扱っているのに対し、ここでは、複文を単位文に分けたものを1つのノードとして扱っている。しか

し、この接続パターンに基づく手法は、対象をマニュアル文に限定し、単位文のタイプは、「ユーザ操作」、「システム可能動作」など、その接続パターンは、「「ユーザ操作」と「ユーザ可能操作」、「ユーザ操作」と「ユーザへの要求」の間には「IF」という関係が成り立つ」などのように、ある程度絞り込むことによって、可能としている。

福本、安原[6]らは、隣接2文間のみならず、文をグループ化し、文とグループ間、2グループ間についても、接続表現や文末表現などの表層的な情報に基づいた手法で意味的な関係の解析を行っている。これは、我々と目的は同じだが、日本語要求仕様文や判例文の解析に重要と思われる条件的、時間的關係についての解析は行われていない。

吉田、横山、西原[8]らは、文中から主題・焦点を抽出した結果を用いて、文脈構造をネットワーク化する方法を提案している。しかし、主題・焦点という文の内容を表す情報を用いているが、導き出されるものは文のつながり方が、「直列型」か「並列型」かといった2文間のつながり方の形だけ、つまり表層的な情報だけである。これに対して我々の目的は2文間の意味的な関係、つまり深層的情報を解析することである。

西澤、中川[9]らは、「だって」、「から」などの接続表現によって、因果関係の原因・理由、および結果の関係が談話中で明示されている場合を対象とし、そのような因果関係が談話中で示す特徴について検討している。我々も因果関係の解析を重要な目的としているが、我々はさらに条件関係や時間的關係などのより広範囲な意味的な関係の解析と目的としている。

1.2. 研究テーマ

以上の現状分析に基づき、本研究では、文間の意味的な関係を表す「文間深層格」を定義し、日本語文章の文と文の間の文間深層格を接続表現などの表層的な情報だけでなく、EDR 電子

・概念体系辞書
約40万の概念に対し、個々に概念IDを与え、それらの間の上下関係を木構造に体系化したものである。

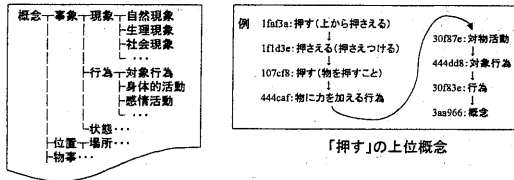


図 1 概念体系辞書 -EDR 電子化辞書-

化辞書[4]による概念 ID などを用いて自動的に決定するシステム InSeRA を開発する。

本論文においては、2章で概念 ID について、3章で文間深層格の定義を、4章で文間深層格の決定方法を、5章でシステム概要を、6章で結論を述べる。

2. 概念 ID -EDR 電子化辞書-

本研究では、文間深層格決定のための重要な情報の1つとして EDR 電子化辞書の概念 ID を用いている。EDR 電子化辞書は言語データを扱った電子化辞書であり、図1のように単語の語意を16進数6桁の概念 ID として与え、それらの間の上下関係を木構造に体系化し、概念体系辞書として集約したものである。

3. 文間深層格

文と文の関係をとらえるには、ある文がどの文につながるか、そしてどのような関係でつながるかを考えなければならない。

本研究では、文間の理由、条件、時間的近接などの意味的な関係を表す21種類の文間深層格を定義した。この文間深層格の詳細な定義を表1にまとめる。これらの文間深層格は表1の第2列目の役割名ごとにその英語表記として定義したものである。これらの役割名の多くは各出典先の参考文献から抽出したものであるが、一部には、「従位接続」を「仮定」、「論理的条件」、「時間的条件」に細分割したように、本研究でさらに細分化を行ったものもある。各深層格の例

日本語文章:例
人は行き先階を持つ

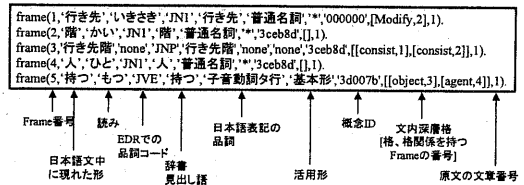


図 2 入力仕様 -意味フレーム-

- 日本語文章:例
1. ジャックは日に1ペニーも貰えない。
 2. あれより早く働かないから。

Ex) interRel (inter-reason, 1, 2).



図 3 出力仕様 -インターリレーション・フレーム-

文については、文番号が () でくくられているものは、各役割の出典先から引用したもので、<>でくくられているものは、独自に作成したものである。なお、この21種類の文間深層格で、オブジェクト指向設計の動的分析や判例文検索における用言間の関係解析には十分であった。

4. 文間深層格の決定方法

4.1. 入力仕様 -意味フレーム-

本システム InSeRA の前段階として本研究室で開発された、1文ごとの意味解析を行う SAGE があり、その出力である意味フレームを InSeRA の入力仕様とする。意味フレームは、図2に示すように10の要素から構成される。

4.2. 出力仕様 -インターリレーション・フレーム-

文と文の間(文番号と文番号の間)にどのような意味的關係があるかを記述した図3に示すようなインターリレーション・フレーム(interrelation frame)を生成する。

表 1 文間深層格の定義

	深層格	役割	詳細な意味	例(〇)のものは各役割の出典先から引用)	格フレームでの表現方法(例と総て)	出典(役割名の出典)
1	inter-reason	理由	Sの理由がSで述べられている	(S1)ドジャックは日に1ベニーも買わない。 (S2)あれより早く働かないから。	interRe(inter-reason, 結果, 理由) interRe(inter-reason, S1, S2)	テキスト言語学入門 自然言語処理
2	inter-cause	原因	Sの原因がSの記述内容となる	(S1)ジャックは倒れた。 (S2)足で踏んで(足から)、頭を怪我した。	interRe(inter-cause, 結果, 原因) interRe(inter-cause, S2, S1)	テキスト言語学入門 自然言語処理
3	inter-enable	実行可能化	Sの事象、状態は、Sの事象、状態を可能にした	(S1)今日、大野君がタルトを作りました。 (S2)南雲がそのタルトを盗んで、すっかり持って行ってしまいました。	interRe(inter-enable, 実行可能になった事象・状態, ある事象を実行可能にした事象・状態) interRe(inter-enable, S2, S1)	テキスト言語学入門
4	inter-purpose	目的	Sの事象、状態は、Sの事象、状態を通じて可能となるように計画されたものである	(S1)おはさんは、戸棚のところへ行きまして。 (S2)可哀相な犬に骨をやらうと思ったからです。	interRe(inter-purpose, 目的のために行う事象・状態, 目的) interRe(inter-purpose, S1, S2)	テキスト言語学入門
5	inter-change	状態変化	Sの状態がSで記述された内容に(通常時間経過に伴い)変化する	(S1)みんな騒いでいた。 (S2)やがて静かになった。	interRe(inter-change, 変化前, 変化後) interRe(inter-change, S1, S2)	自然言語処理
6	inter-detail	詳細化	Sで述べられた事象、状態、またその要素についての詳しい内容がSで述べられている	(S1)そこにあったのは、義務教育段階の子供についてだけである。 (S2)つまり、小学生と中学生についてだけである。	interRe(inter-detail, 本文, 本文についてのさらに詳しい内容) interRe(inter-detail, S1, S2)	自然言語処理
7	inter-example	例提示	Sで述べられた事象、状態の具体例の項目がSで提示される	(S1)それぞれの立場によって状況は違ってくる。 (S2)例えば公務員の場合、時の総理大臣を公に批判するなど考えられることはない。	interRe(inter-example, 本文, 本文の具体例の項目) interRe(inter-example, S1, S2)	自然言語処理
8	inter-explanation	例説明	Sで述べられた事象、状態の具体例の説明がSで行われる	(S1)例えば公務員の場合、時の総理大臣を公に批判するなど考えられることはない。 (S2)つまり、自分の身内の批判をするわけにはいかないということである。	interRe(inter-explanation, 本文, 本文で述べられた具体例の説明) interRe(inter-explanation, S1, S2)	自然言語処理
9	inter-answer	質問-応答	Sの質問に対してSで答えが示される	(S1)天文学者はこの種の衝突になぜ興味を持つのだろうか。 (S2)その答えは、“熱”を発生させるのに連星が演じている役割にある。	interRe(inter-answer, 質問, 応答) interRe(inter-answer, S1, S2)	自然言語処理
10	inter-contrast	対比	SとSが対比関係にある事象や状態などの記述である	(S1)彼女は一所懸命勉強している。 (S2)それと比べ、俺は怠けている。	interRe(inter-contrast, 前文, 後文) interRe(inter-contrast, S1, S2)	自然言語処理
11	inter-contradiction	逆接	SとSが同じ地位にあるけど、一見互いに矛盾し、両立しないように見える	(S1)彼らは、これは放逐になりそうだと言っていた。 (S2)かま。最後の最後になって外交的勝利をやってのけた。	interRe(inter-contradiction, 前文, 後文) interRe(inter-contradiction, S1, S2)	テキスト言語学入門
12	inter-or	選言	SとSが択一的な地位である	(S1)A氏が、その候補に立腹しているのならば話は別ですが、もしそうなら、 (S2)恐ろしいほどしかたらず、A氏の介添人が侮蔑されると感じているということもありえますね。	interRe(inter-or, 前文, 後文) interRe(inter-or, S1, S2)	テキスト言語学入門
13	inter-logicalCondition	論理的条件	SとSが条件関係 SがSの事象・状態が起こるための条件となっている	(S1)AとBが等しい。 (S2)そうならば、Cを出力する。	interRe(inter-logicalCondition, 条件が真のとき起こる事象・状態, 条件) interRe(inter-logicalCondition, S2, S1)	テキスト言語学入門から抽出した「従位接続」を細分化
14	inter-timingCondition	時間的条件	SとSが条件関係 SがSの事象・状態が起こるきっかけになっている	(S1)エレベーターが着いた。 (S2)すると降りる人が中から出てきた。	interRe(inter-timingCondition, 条件・きっかけによって起こる事象・状態, 条件・きっかけ) interRe(inter-logicalCondition, S2, S1)	テキスト言語学入門から抽出した「従位接続」を細分化
15	inter-subjunctive	仮定	SとSが仮定とその結果起こるかもしれない、あるいは起こったかもしれない、どのような想像上の事象・状態を表す。	(S1)その試合には勝てなかっただろう。 (S2)もっと練習したとて立。	interRe(inter-subjunctive, 仮定によって予想される結果, 仮定) interRe(inter-subjunctive, S1, S2)	テキスト言語学入門から抽出した「従位接続」を細分化
16	inter-sequence	時間的順接	SとSが表す事象・状態が互いに時間的に順序付けられている	(S1)大統領は、「帰国してうれい」と感動した様子で述べた。 (S2)それから、彼らが開きにきた話をした。	interRe(inter-sequence, 時間的に前に起きた事象・状態, 時間的に後に起きた事象・状態) interRe(inter-sequence, S1, S2)	テキスト言語学入門から抽出した「時間的順接」を細分化
17	inter-cooccurrence	時間的重複	SとSが表す事象・状態が時間的に重複している	(S1)翌日、A国内閣も満場一致で協定の最終的細目を可決した。 (S2)その間に、B国とA国の国防相がワシントンで会談を行った。	interRe(inter-cooccurrence, 本文, 本文と同時に起きていた事象・状態) interRe(inter-cooccurrence, S1, S2)	テキスト言語学入門から抽出した「時間的順接」を細分化
18	inter-synchronous	時間的同時開始	SとSが表す事象・状態は同時に開始されている	(S1)ピッチャーが球を投げた。 (S2)それと同時に、ランナーが走り出した。	interRe(inter-synchronous, 本文, 本文と同時に起きた事象・状態) interRe(inter-synchronous, S1, S2)	テキスト言語学入門から抽出した「時間的順接」を細分化
19	inter-equivalent	等位	SとSが同じ地位である	(S1)大きな鳥たちは、町のすぐ外にある公園の木々を好んでねぐらにしていた。 (S2)1985年以降、地元住民はその状況を最大限利用してきた。	interRe(inter-equivalent, 前文, 後文) interRe(inter-equivalent, S1, S2)	テキスト言語学入門
20	inter-conversion	転換	SとSの間に何の関係もなく、話題が転換している	(S1)私はお墨御飯を食った。 (S2)ところで、私には構えるべきものがある。	interRe(inter-conversion, 前文, 後文) interRe(inter-conversion, S1, S2)	現代日本語文法入門
21	inter-parallel	並列	SとSの関係と、SとSの関係が同一または同様の関係であり、SとSが並列関係にある	(S0)その答えは、“熱”を発生させるのに連星が演じている役割にある。 (S1)連星と単星が衝突する際、連星は…、その周囲の星の集団を暖めることがある。 (S2)また、意図によって連星の軌道が縮小し、そのために高密度の星団の中核の温度が上昇することも考えられる。	interRe(inter-parallel, 前文, 後文) interRe(inter-parallel, S1, S2)	自然言語処理

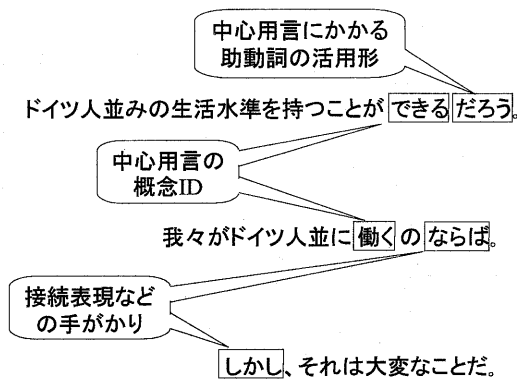


図 4 文間深層格決定に用いる情報

このフレームは、第 1 引数で示す文間深層格名で表した意味的關係が、第 2 引数の文から、第 3 引数の文に向けて存在することを表している。

4.3. 文間深層格の決定方法

隣接 2 文間に我々が提案するヒューリスティックなルールを適用することによって、2 文間に文間深層格を割り当てる。本研究では、文間深層格を決定するには、図 4 に示すような接続語などの表層的な手がかり、中心用言の概念 ID、助詞の活用形などを用いる。文間深層格を決定するためのこれらのルールは大きく以下の 3 つに分類できる。

① 接続表現と概念 ID による方法

例えば、「理由」「原因」「目的」のように同一の接続表現を持つものに対して、その中心用言の上位概念の違いによって、適切な文間深層格を割り当てる。

② 接続表現と助動詞の活用形による方法。

「仮定」「論理的条件」のように同一の接続表現を持つものに対して、その中心用言にかかる助動詞の活用形の違いによって、適切な文間深層格を割り当てる。

③ 接続表現による直接的対応付け

①や②とは異なり、単に接続表現のみで文間深層格を決定する。例えば、「その間」という接続表現を持つ文とその前文との間

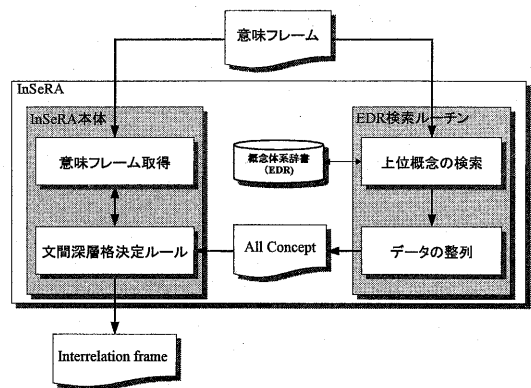


図 5 InSeRA の処理の流れ

に、「時間的重複」という文間深層格を割り当てる。

4.4. 各文間深層格の決定ルール

文間深層格ごとにその決定ルールを表 2 にまとめた。この表は、当該文 S が接続表現の列に指定した接続表現を持ち、かつ決定ルールの列に指定された条件を満たすとき、出力フレームの列に指定した書式で文間深層格を出力することを表している。ただし一部接続表現を用いないものも存在する。

なお、表中の決定ルールの列で、「条件なし」とあるのは、前節の分類③に属するもので、その他は①か②であり、その詳細は表中に記述されている。また、これらのルールの中で接続表現と文間深層格の対応については、一部は本研究で分類したものであるが、他は以下の出典先から採用したものである。これらについては、表 2 の第 4 列にその出典先を示した。

*1: 本研究で分類したもの

*2: 小池清治; 現代日本語文法; p244

*3: 池上嘉彦 他 共訳; テキスト言語学入門; 第 4 章—結束構造

5. システム概要

本システム InSeRA は図 5 のように InSeRA 本体と EDR 検索ルーチンからなる。

表 2 文間深層格決定ルール

文間深層格	役割名	決定ルール 一般的には、当該文Sが右列に指定した接続現を持ち、かつ以下の条件を満たす。	接続表現(*の後は出典番号)	出力フレーム	
1	inter-reason	理由	順1:Sが右列の順1に指定した接続表現を持ち、Sの中心用言の上位概念が「行為」(30f83e)。ただし、「思考する」(444dda)と「望む」(30f87d)はのぞく。 順2:Sが右列の順2に指定した接続表現を持ち、Sの前置の中心用言の上位概念が「行為」(30f83e)。ただし、「思考する」(444dda)と「望む」(30f87d)はのぞく。	{だから、なので、故に、というの} *3、{～から、～だから、なぜなら} *1、{～から、というの、なぜなら} 順1 だから、なので、故に、それ故} 順2	順1: interRe(inter-reason, Sの前置, S) 順2: interRe(inter-reason, S, Sの前置)
2	inter-cause	原因	順1:Sが右列の順1に指定した接続表現を持ち、Sの中心用言の上位概念が「現象」(30f7e5)。 順2:Sが右列の順2に指定した接続表現を持ち、Sの前置の中心用言の上位概念が「現象」(30f7e5)。	{だから、なので、故に、というの} *3、{そして、～から、～だから、なぜなら} *1、{～から、というの、なぜなら} 順1 だから、なので、故に、それ故、そして} 順2	順1: interRe(inter-cause, Sの前置, S) 順2: interRe(inter-cause, S, Sの前置)
3	inter-enable	実行可能化	接続表現は使わない。 当該文Sの中心用言の上位概念が「創造する」(30f89e)。		interRe(inter-enable, Sの前置, S)
4	inter-purpose	目的	Sの中心用言の上位概念が「思考する」(444dda)、あるいは「望む」(30f87d)。	{～から} *1	interRe(inter-purpose, Sの前置, S)
5	inter-change	状態変化	Sの中心用言の上位概念が「変化」(30f85e)。	{やがて、そして} *1	interRe(inter-change, Sの前置, S)
6	inter-detail	詳細化	Sの前置に例提示を表す接続表現が含まれていない。	{すなわち、つまり、要するに} *2、	interRe(inter-detail, Sの前置, S)
7	inter-example	例提示	条件なし。(接続表現の存在による直接的対応付け)。	{例えば} *2、	interRe(inter-example, Sの前置, S)
8	inter-explanation	例説明	Sの前置に例提示を表す接続表現が含まれている。	{すなわち、つまり、要するに} *2、	interRe(inter-explanation, Sの前置, S)
9	inter-answer	質問-応答	質問文を表す接続表現を含む文を質問とし、応答文を表す接続表現を含む文を応答とする。 質問文を表す接続表現を含む文を質問とし、その後文を応答とする。	質問文:{～か、～だろう(か)、～でしょうか} *1 応答文:{(その)答え(は)、それ(は)} *1、	interRe(inter-answer, 質問, 応答)
10	inter-contrast	対比	条件なし。(接続表現の存在による直接的対応付け)。	{それに比べ、一方～は、それに対して} *1、	interRe(inter-contrast, Sの前置, S)
11	inter-contradiction	逆接	条件なし。(接続表現の存在による直接的対応付け)。	{けれども、しかし、しかしながら、しかるに、そのくせ、それでも、それなのに、それいしても、それにもかかわらず、だが、ただ、なのに、でも、とはいえ、とはいうものの} *2、{しかし、しかしながら、それでも、それにもかかわらず} *3、	interRe(inter-contradiction, Sの前置, S)
12	inter-or	選言	条件なし。(接続表現の存在による直接的対応付け)。	{あるいは、または、もしくは} *3、	interRe(inter-or, Sの前置, S)
13	inter-logicalCondition	論理的条件	Sの前置の中心用言に活用形が「推量形」の助動詞がかかってない。	{なら(ば)} *1、	interRe(inter-logicalCondition, Sの前置, S)
14	inter-timingCondition	時間的条件	条件なし。(接続表現の存在による直接的対応付け)。	{すると} *1、	interRe(inter-timingCondition, S, Sの前置)
15	inter-subjunctive	仮定	Sの前置の中心用言にかかる助動詞の活用形が「推量形」。	{～としても、～だとしても、～なら(ば)} *3、	interRe(inter-subjunctive, Sの前置, S)
16	inter-sequence	時間的順接	条件なし。(接続表現の存在による直接的対応付け)。 注)接続表現によって、順序が違う。	{(そしたら) *2、(それから、次に、～の前に、～の後に、～以來) *3、 注1: 「すると」は、時間的順接から削除。時間的条件だけとする。 (そしたら、それから、次に、～の後に、～以來) 順1 (～の前に) 順2	順1: interRe(inter-sequence, Sの前置, S) 順2: interRe(inter-sequence, S, Sの前置)
17	inter-cooccurrence	時間的重複	条件なし。(接続表現の存在による直接的対応付け)。	{(その間、その時) *3、(その際) *1、	interRe(inter-cooccurrence, Sの前置, S)
18	inter-synchronous	時間的同時開始	条件なし。(接続表現の存在による直接的対応付け)。	{(それ)と同時に} *1、	interRe(inter-synchronous, Sの前置, S)
19	inter-equivalent	等位	条件なし。(接続表現の存在による直接的対応付け)。	{したがって、そして、そこで、それで、それゆえ、それなら、故に} *2 {そして、さらに、その上} *3、	interRe(inter-equivalent, Sの前置, S)
20	inter-conversion	転換	条件なし。(接続表現の存在による直接的対応付け)。	{(さて、では、それでは、と併に) *2、(ところで) *1、	interRe(inter-conversion, Sの前置, S)
21	inter-parallel	並列	条件なし。(接続表現の存在による直接的対応付け)。	{また、その上、それに、しかも、なお} *1	interRe(inter-parallel, Sの前置, S)

さらに、InSeRA 本体は意味フレーム取得部と文間深層格決定ルールからなり、EDR 検索ルーチンは上位概念検索部とデータ整理部からなる。

入力された意味フレームは InSeRA 本体と EDR 検索ルーチンの両方に渡される。EDR 検索ルーチンでは、意味フレーム中の語の上位概念を最上位まですべて検索し、その結果を整理し、重複データを削除し All Concept ファイルとして InSeRA 本体に引き渡す。InSeRA 本体は、

意味フレーム取得部で該当文の文番号 S_i の意味フレーム群とその前後の文 $S_{i\pm 1}$ の意味フレーム群から必要な情報を抽出し、ルール部でこの意味フレーム群の間に抽出された情報と All Concept ファイルにまとめられた上位概念の情報を基に表 2 にまとめたヒューリスティックなルールを適用することによって文間深層格を割り当てる。この結果をインターリレーション・フレームとして出力する。図 6 に prolog で記述されたルールの一部を示す。

7. まとめ

7.1. 結論

本研究では、文間の意味的な関係を表す「文間深層格」を定義し、日本語文章の文間深層格を接続表現などの表層的な情報だけでなく、EDR 電子化辞書による概念 ID などを用いて自動的に決定するシステム InSeRA を開発した。

これによって、機械的に文間の深層格を求めることができるようになった。また、「理由」、「原因」、「目的」のように接続表現だけでは細分化できなかった文間関係を解析することに成功した。

7.2. 今後の課題

今回の本システムの評価は、エレベータ問題などの本研究室で作成した例文や長尾氏、角田氏らがまとめた「自然言語処理」、ボウグランド (R.de.Beaugrande) とドレスラー (W.U.Dressler) の両氏の著書「テキスト言語学入門」から採用したわずかな事例を基に行った。これらに関しては、正しい結果が得られた。

しかし、より客観的で信頼できる評価を行うには、より大量のデータを基に実験を行う必要がある。ところが、文間関係には、意味解析におけるコーパス辞書のように正解事例集のようなものが存在しない。今後はコーパス辞書のような正解事例集を第3者に依頼して作成してもらう必要がある。

一方、文脈解析全体を自動化するという目的からすれば、主題・焦点の解析や照応解析など残された問題を解決する必要がある。

【参考文献】

- [1] 長尾真, 佐藤理史, 黒橋禎夫, 角田達彦: 自然言語処理, 岩波出版 (1996).
- [2] R.de. ボウグランド, W. ドレスラー: テキスト言語学入門, 紀伊国屋書店 (1984).
- [3] 小池清治: 現代日本語文法入門, 筑摩書房 (1998).

- [4] (株) 日本語電子辞書研究所: EDR 電子化辞書使用説明書 (第2版), (株) 日本語電子辞書研究所 (1995).
- [5] 原田実, 水野高宏: EDR を用いた日本語意味解析システム SAGE, 人工知能学会論文誌, Vol.16, No.1, pp.85-93 (2001.1).
- [6] 福本淳一, 安原宏: 文の接続関係解析に基づく文章構造解析, 情報処理学会報告, 情報処理学会, 92-NL-88, pp.9-16 (1992).
- [7] 田中智博, 林良彦: 文の連携パターンに基づく日本語テキスト構造の解析, 情報処理学会報告, 情報処理学会, 93-NL-89, pp.17-24 (1992).
- [8] 吉田悦子, 横山晶一, 西原典孝: 主題間の関係を用いた文脈構造ネットワークの構築, 情報処理学会報告, 情報処理学会, 98-NL-124, pp.17-24 (1998).
- [9] 西澤信一郎, 中川裕志: 日本語の自由会話における談話構造の推定 ~因果関係を表す接続詞の場合~, 自然言語処理, 言語処理学会, vol.4, No.4, pp.61-72 (1997).
- [10] 原田実, 矢後友和: 日本語要求仕様文章からオブジェクト指向による動的モデルを生成する CAMEO/D の開発, 情報処理学会第62回全国大会論文集, 印刷中 (2001.3).