

## 日本語時間表現の一解釈法

溝 渕 昭 二<sup>†</sup> 住 友 徹<sup>†</sup>  
 弘 田 正 雄<sup>††</sup> 青 江 順 一<sup>††</sup>

自然言語による時間表現に対して、時間、時点、区間などの意味を解釈し、さらに年、月などの属性とその数量を考慮して形式表現を決定することは、自然言語インタフェースにおける問合せ文などの理解において、非常に重要な課題である。この課題を解決するために、本論文では時間表現の意味解釈を時点、時点区間などの6種類の対象概念に、さらにそれを15種類の下位対象概念に分類する。また、対象概念に開始、周期などの補足的な意味情報を与える補足概念を10種類に分類する。これらの分類に従って、時間表現に対応する形式表現を定義し、形態素列からなる時間表現を形式表現に変換するアルゴリズムを提案する。このアルゴリズムは、構文解析におけるshift-reduce法を改良したものである。提案手法により、多彩な自然言語の時間表現は意味と数量的な情報を加味した正確な式で表現できるので、その利用価値は高いものとなる。新聞、雑誌、問合せ文などの約2,000の時間表現データに対して、本手法を適用し、正しい解釈結果が93%以上になることを実証する。

## A Method for Understanding Japanese Time Expressions

SHOJI MIZOBUCHI,<sup>†</sup> TORU SUMITOMO,<sup>†</sup> MASAO FUKETA<sup>††</sup>  
 and JUN-ICHI AOE<sup>††</sup>

In natural language processing systems, it is very important to understand the meaning of time expressions and convert them into formal representations which include quantitative values such as the number of years and months, period lengths and so on. This paper proposes a method for improving the understanding of time expressions consisting of noun phrases. First, the meaning of an expression is classified into six categories for which a formal representation is defined. The proposed representation includes both conceptual and quantitative aspects in comparison with traditional approaches such as those of machine translation systems, and allows for more precise understanding of time expressions. An algorithm for translating time expressions into the corresponding formal representation by using a shift-reduce parsing technique is presented. Experiments for about 2,000 time expressions extracted from actual documents show that the proposed method yields correct interpretation in about 93% of the cases.

## 1. ま え が き

自然言語インタフェース<sup>1)~3)</sup>における問合せ文の理解において、時間表現の意味を正確に理解することは非常に重要な課題である。また、文書の要約システム<sup>1),4),5)</sup>において数量を含めた時間表現のより深い意味を的確に把握できれば、それを視覚的な要約表現で表すことも可能になり、自然言語処理システムにおいて、本研究内容は意義深いものとなる。

機械翻訳<sup>1),6),7)</sup>でも時間表現の解析研究は行われているが、自然言語における時間表現は非常に多彩であり、その解析には多くの解析規則を必要とする。しかも、機械翻訳での解釈は、あくまで表層表現を翻訳(昨日を yesterday と訳す)するのが目的であって、時間表現を年、月、日などの値として正確に解釈し、その意味(時間、時点、区間など)を捉えることを目的としていない。これに対して、より高度な自然言語インタフェースや視覚的要約処理が必要としているのは、多彩な時間表現に対するより詳細で正確な意味情報の獲得である。たとえば、今日が1997年11月12日であれば、「明日」の意味は、「今日の次の日」という解釈だけではなく、1997年11月13日という数量的な解釈もともなうものである。

以上の課題に対して、本論文では、時間表現の意味

<sup>†</sup> 徳島大学大学院工学研究科  
 Graduate School of Engineering, University of  
 Tokushima

<sup>††</sup> 徳島大学工学部知能情報工学科  
 Department of Information Science and Intelligent Sys-  
 tems, University of Tokushima

解釈を時点、時点区間などの6種類の概念(対象概念と呼ぶ)に、さらに対象概念を15種類の下位対象概念に分類する。また、対象概念の解釈に開始、周期などの補足的な意味情報を与える概念(補足概念と呼ぶ)を10種類に分類する。これらの分類に対して、時間表現に対応する形式表現を定義し、形態素列からなる時間表現を形式表現に変換するアルゴリズムを提案する。提案手法により、多彩な自然言語の時間表現はその意味と数量的な情報を加味した正確な式で表現できるので、その利用価値はきわめて高いものとなる。

以下、2章では時間表現の意味解釈の分類を説明し、3章では、これら分類に対する表現形式を定義する。4章では、時間表現から表現形式に変換するアルゴリズムを提案し、5章で実際の自然言語の時間表現に対する実験より、提案手法を評価する。最後に、6章では本論文のまとめと、今後の課題について触れる。

2. 時間表現に対する概念と分類

時間表現には、名詞句だけでなく、テンスやアスペクトの表現<sup>(8),(9)</sup>も存在するが、本論文では数量情報を含む名詞句の時間表現を議論の対象とする。また、以下で述べる概念の解釈や分類には、文献8), 10)~13)を参考にした。

2.1 時間表現に対応する概念

時間情報を表す形態素列を成分と呼び、そのうち分割すると時間情報を失ってしまうような成分の最小単位を基本成分と呼ぶ。以後、成分は「」で囲んで示す。基本成分を除く成分はそれより小さな成分から構成されるので、時間表現は図1(a)のような木構造で表すことができる。たとえば、「今日から」は、基本成分である「今日」と「から」で構成される。

成分が表す時間の意味を概念と呼び、これにより成分間の構成関係が決定する。以後、概念は《》で囲んで示す。概念は成分と対応しているのので、時間表現に対応する概念も図1(b)のような木構造で表すことができる。たとえば、「今日から」の概念《開始時点》は、「今日」の概念《時点》と「から」の概念《開始》から構成される。

成分と概念の対応関係は一意ではない。成分には複数の概念を持つものがあり、そのうちのどれが成分の概念となるかは他の成分の概念との関係によって決定される。したがって、表記が同じでも概念の種類によっては、接続できたりできなかったりする成分が存在する。

2.2 概念の分類

概念は、「今日」や「1998年」のように時間的な対

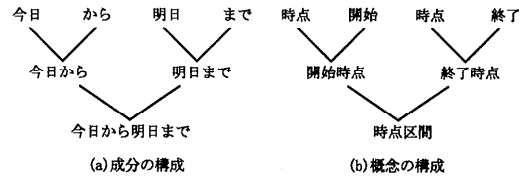


図1 成分と概念の構成例

Fig. 1 An example of constitution of components and concepts.

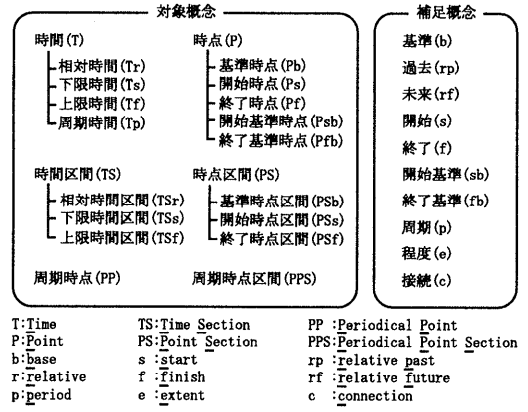


図2 概念の分類

Fig. 2 Classification of concepts.

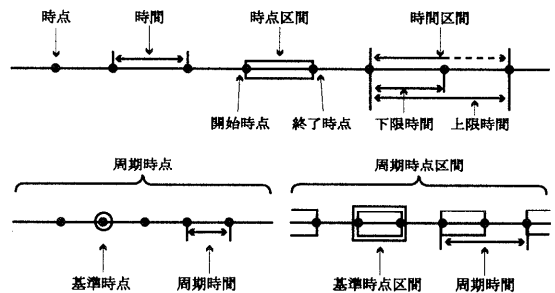


図3 対象概念の視覚化

Fig. 3 Visualization of concepts.

象を表す対象概念と、「から」や「より」のように補足的な情報を表す補足概念とに分類される。図1(b)では《開始》と《終了》が補足概念で、それ以外は対象概念である。対象概念と補足概念の分類を図2に定義する。概念名に付く括弧内の記号は略称を表す。また、理解を助けるために、視覚化した対象概念を図3に示す。

2.2.1 対象概念

- 《時間》は時の長さを表し、年、月、日などの数量を値を持つ。成分には、「半年(6ヵ月)☆」,「1年

☆ 成分「」内の()の情報は理解を助けるために付けたものである。以後同様。

3ヶ月],「365日」などがある。

- 《時点》は時の流れの中の1点を表し、《時間》と同様に、年、月、日などの数量を値に持つ。成分には、「1997年12月10日」や「昭和47年4月」のような暦の表現、「10年後」や「今日から25日後」のようなある時点からの時間距離の表現がある。
- 《時間区間》は不特定な時間を表し、《下限時間》と《上限時間》から構成される。成分には、「1年から2年」のような時間の下限と上限を両方含む表現、「半年以上」や「半年以下」のような時間の下限と上限のどちらか1つだけを含む表現、「約10日間」のような時間と程度の組合せ表現がある。「1年から2年」と「約10日間」をさらに小さな成分に分解すると、「1年から」、「から2年」、「約」、「10日間」となる。これらに対応する概念はそれぞれ、《下限時間》、《上限時間》、《程度》、《時間》である。
- 《時点区間》は2つの時点に挟まれた区間であり、《開始時点》と《終了時点》から構成される。成分には、「春(3月から5月まで)」のような開始と終了の時点を両方含む表現、「今年から」や「今年まで」のような開始と終了の時点をどちらか1つだけを含む表現、「今年から2年間」のような開始と終了のうちどちらかの時点と時間の組合せ表現、「6月頃」のような時点と程度の組合せ表現、「今年から数年後」のようなある時点からの不特定な時間距離の表現がある。「今日から2日間」と「今年から数年後」の成分をさらに小さな成分に分解すると、「今日から」、「2日間」、「今年から」、「数年後」となる。これらに対応する概念はそれぞれ、《開始基準時点》、《時間》、《基準時点》、《相対時間区間》である。
- 《周期時点》はある1点の基準から等間隔の位置に存在する時点であり、《基準時点》と《周期時間》から構成される。成分には、「4年ごと」や「毎年」などがある。
- 《周期時点区間》は、ある1区間の基準から等間隔の位置に存在する時点区間であり、《基準時点区間》と《周期時間》から構成される。成分には、「毎年春」や「毎月初旬」などがある。
- 《時間》の性質を継承する下位概念には、《相対時間》、《下限時間》、《上限時間》、《周期時間》があり、それぞれ「半年後」、「半年以上」、「半年以下」、「半年ごと」などの成分が対応する。
- 《時点》の性質を継承する下位概念には、《基準時点》、《開始時点》、《終了時点》、《開始基準時点》、《終

了基準時点》があり、それぞれ「今日から(25日後)」、「今日から(明日まで)」、「(今日から)明日まで」、「今日から(3日間)」、「明後日まで(2日間)」などの成分が対応する。

- 《時間区間》の性質を継承する下位概念には、《相対時間区間》、《下限時間区間》、《上限時間区間》があり、それぞれ「約半年後」、「約半年以上」、「約半年以下」などの成分が対応する。
- 《時点区間》の性質を継承する下位概念には、《基準時点区間》、《開始時点区間》、《終了時点区間》があり、それぞれ「(毎年)春」、「春から(夏まで)」、「(春から)夏まで」などの成分が対応する。

### 2.2.2 補足概念

- 《基準》は付属する《時点》、あるいは《時点区間》が時間を計測するための基準情報になることを示す。成分には、「(今日)から(2日後)」などがある。
- 《未来》は時の流れる方向を示し、《過去》はその逆方向を示す。成分には、それぞれ「(2日)後」、「(2日)前」などがある。
- 《開始》は始まりを、《終了》は終わりを示す。成分には、それぞれ「(今日)から」、「(明日)まで」などがある。
- 《周期》は繰返しを示す。成分には、「(2日)ごと」、「(2日)おき」などがある。
- 《開始基準》は《開始》と《基準》を同時に示し、《終了基準》は《終了》と《基準》を同時に示す。成分には、それぞれ「(今日)から(2日間)」、「(明日)まで(2日間)」などがある。
- 《程度》は数量に幅があることを示す。成分には、「約」、「程」などがある。
- 《接続》は時間表現の接続を示す。「1998年の3月」における「の」は、「1998年」と「3月」を「1998年3月」に接続する。

## 3. 概念の表現形式と構成関係

### 3.1 概念名と属性値による表現形式

時間表現の意味である概念名  $c^*$  とその属性値列  $v$  により、時間表現の解釈結果を形式的に  $[c:v]$  で表し、これをフレームと呼ぶ。この表現は、知識表現であるフレーム<sup>14),15)</sup>からフレーム名とスロット値だけを抽出したものである $\star\star$ 。したがって、属性値がフレーム

$\star$  図2の略称で記す。

$\star\star$  本論文で定義するフレームのスロット名とその並びは、あらかじめ年、月、日などに限定できるので、形式表現ではスロット名を省略する。

になる場合もある。ただし、属性値列  $v$  は、 $c$  が補足概念のとき定義されないので、 $[c:]$  と記述する。以下に、対象概念に対するフレームを定義する。

(1) 《時間》(T)

属性である年、月、日の値をそれぞれ  $y_t, m_t, d_t$  とするとき $\star$ 、フレームは次で表される。

$$[T:y_t, m_t, d_t]$$

$y_t, m_t, d_t$  は次のいずれかの条件を満足する整数であり、無限大の値は  $\infty$  で表す。成分から値が解釈できない場合は、属性値を 0 で表す。

$$0 \leq y_t, 0 \leq m_t \leq 11, 0 \leq d_t \quad (1)$$

$$y_t \leq 0, -11 \leq m_t \leq 0, d_t \leq 0 \quad (2)$$

たとえば、「1年3ヶ月」、「365日」に対応するフレームはそれぞれ、 $[T:1, 3, 0]$ 、 $[T:0, 0, 365]$  である。

属性値の範囲は、次式により適宜変更する。

$$[T:y_t, m_t, d_t] = [T:y_t + 1, m_t - 12, d_t] \quad (3)$$

$$[T:y_t, m_t, d_t] = [T:y_t - 1, m_t + 12, d_t] \quad (4)$$

たとえば、 $[T:0, 15, 0]$  の場合、式 (3) を使用して属性値が条件 (1) を満たすように変更する。

$$[T:0, 15, 0] = [T:0 + 1, 15 - 12, 0] = [T:1, 3, 0]$$

また、 $t = [T:y_t, m_t, d_t]$  とするとき、次の演算を定義する。

$$-t = [T: -y_t, -m_t, -d_t] \quad (5)$$

たとえば、 $t = [T:1, 3, 0]$  とするとき、式 (5) の演算結果を次に示す。

$$-t = [T: -1, -3, 0]$$

(2) 《時点》(P)

属性である年、月、日の値をそれぞれ  $y_p, m_p, d_p$  とするとき、フレームは次で表される。

$$[P:y_p, m_p, d_p]$$

$y_p, m_p, d_p$  は、次の条件を満足する整数である。成分から値が解釈できない場合は、 $\epsilon$  で表す $\star\star$ 。

$$1 \leq y_p, 1 \leq m_p \leq 12, 1 \leq d_p \leq D(y_p, m_p) \quad (6)$$

ここで、 $D(y_p, m_p)$  は次に定義する関数である。

$$D(y_p, m_p) = \begin{cases} 28 & y_p \neq \text{閏年かつ } m_p = 2 \\ 29 & y_p = \text{閏年かつ } m_p = 2 \\ 30 & m_p = 4, 6, 9, 11 \\ 31 & m_p = 1, 3, 5, 7, 8, 10, 12 \end{cases}$$

たとえば、「1997年12月10日」のフレームは  $[P:1997, 12, 10]$  で表される。

属性値の範囲は次式により適宜変更する。

$$[P:y_p, m_p, d_p]$$

$\star$  時、分、秒などの属性も存在するが、記述を簡単にするために年、月、日で説明する。

$\star\star$  《時間》では値が決定できない場合は 0 でよいが、《時点》では未定義であるので、 $\epsilon$  と記述する。

$$= [P:y_p + 1, m_p - 12, d_p] \quad (7)$$

$$[P:y_p, m_p, d_p] = [P:y_p - 1, m_p + 12, d_p] \quad (8)$$

$$[P:y_p, m_p, d_p] = [P:y_p, m_p + 1, d_p - D(y_p, m_p)] \quad (9)$$

$$[P:y_p, m_p, d_p] = [P:y_p, m_p - 1, d_p + D(y_p, m_p - 1)] \quad (10)$$

たとえば、 $[P:1997, 12, -15]$  の場合、式 (10) を使用して属性値が条件 (6) を満たすように変更する。

$$[P:1997, 12, -15] = [P:1997, 12 - 1, -15 + 30]$$

$$= [P:1997, 11, 15]$$

(3) 《時間区間》(TS)

属性である開始時間、終了時間の値をそれぞれ、 $t_s, t_f$  とするとき、フレームは次で表される。

$$[TS:t_s, t_f]$$

$t_s, t_f$  は《時間》、あるいは《時間区間》のフレームである。たとえば、「1年から2年」、「約2年以上」のフレームはそれぞれ次で表される。ただし、後者は程度「約」の幅を1年とした場合の結果である。

$$[TS:[T:1, 0, 0], [T:2, 0, 0]]$$

$$[TS:[TS:[T:1, 0, 0], [T:3, 0, 0]], [T:\infty, \infty, \infty]]$$

$ts = [TS:t_s, t_f]$  とするとき、次の演算を定義する。

$$-ts = [TS: -t_f, -t_s] \quad (11)$$

たとえば、 $ts = [TS:[T:0, 0, 2], [T:0, 0, 3]]$  とするとき、式 (11) の演算結果を次に示す。

$$-ts = [TS: -[T:0, 0, 3], -[T:0, 0, 2]] = [TS:[T:0, 0, -3], [T:0, 0, -2]]$$

(4) 《時点区間》(PS)

属性である開始時点、終了時点の値をそれぞれ、 $p_s, p_f$  とするとき、フレームは次で表される。

$$[PS:p_s, p_f]$$

$p_s, p_f$  は《時点》、あるいは《時点区間》のフレームである。たとえば、「今年(1998年)の春」、「1世紀から」のフレームはそれぞれ次で表される。

$$[PS:[P:1998, 3, \epsilon], [P:1998, 5, \epsilon]]$$

$$[PS: [PS:[P:1, \epsilon, \epsilon], [P:100, \epsilon, \epsilon]], [P:\infty, \infty, \infty]]$$

(5) 《周期時点》(PP)

属性である基準時点、周期時間の値をそれぞれ、 $p_b, t_p$  とするとき、フレームは次で表される。

$$[PP:p_b, t_p]$$

$p_b, t_p$  はそれぞれ《時点》と《時間》のフレームである。たとえば、「(1997年を基準として)4年ごと」のフレームは、 $[PP:[P:1997, \epsilon, \epsilon], [T:4, 0, 0]]$  となる。

(6) 《周期時点区間》(PPS)

属性である基準時点区間、周期時間の値をそれぞれ、

$ps_b, t_p$  とするとき、フレームは次で表される。

$$[PPS:ps_b, t_p]$$

$ps_b, t_p$  はそれぞれ《時点区間》と《時間》のフレームである。たとえば、「毎年春 (=1997年春を基準にして1年ごと)」に対応するフレームは、 $[PPS:[PS:[P:1997,3,\epsilon], [P:1997,5,\epsilon]], [T:1,0,0]]$  となる。「毎年」を「1年ごと」として解釈する場合、どの年でもその基準となりうるので、《時点》のフレームの年属性は1997でなくてもよい。この値を選択した理由は、この表現の発話された時点が1997年だからである。

#### (7) 《時間》の下位概念 (**Tr, Ts, Tf, Tp**)

《相対時間》, 《下限時間》, 《上限時間》, 《周期時間》の属性である時間の値を  $t$  とするとき、これらのフレームはそれぞれ、

$$[Tr:t], [Ts:t], [Tf:t], [Tp:t]$$

で表される。 $t$  は《時間》のフレームである。たとえば、《相対時間》を表す「半年後」に対応するフレームは  $[Tr:[T:0,6,0]]$  である。

#### (8) 《時点》の下位概念 (**Pb, Ps, Pf, Psb, Pfb**)

《基準時点》, 《開始時点》, 《終了時点》, 《開始基準時点》, 《終了基準時点》の属性である時点  $p$  とするとき、これらの概念はそれぞれ、

$$[Pb:p], [Ps:p], [Pf:p], [Psb:p], [Pfb:p]$$

で表される。 $p$  は《時点》のフレームである。たとえば、《基準時点》を表す「1998年から」に対応するフレームは  $[Pb:[P:1998,\epsilon,\epsilon]]$  である。

#### (9) 《時間区間》の下位概念 (**TSr, TSs, TSf**)

《相対時間区間》, 《開始時間区間》, 《終了時間区間》の属性である時間区間の値を  $ts$  とするとき、これらの概念はそれぞれ、

$$[TSr:ts], [TSs:ts], [TSf:ts]$$

で表される。 $ts$  は《時間区間》のフレームである。たとえば、《相対時間区間》を表す「約半年後」に対応するフレームは  $[TSr:[TS:[Pb:0,5,0], [Pf:0,7,0]]]$  である。

#### (10) 《時点区間》の下位概念 (**PSb, PSs, PSf**)

《基準時点区間》, 《開始時点区間》, 《終了時点区間》の属性である時点区間の値を  $ps$  とするとき、これらの概念はそれぞれ、

$$[PSb:ps], [PSs:ps], [PSf:ps]$$

で表される。 $ps$  は《時点区間》のフレームである。たとえば、《開始時点区間》を表す「1998年春から」に対応するフレームは  $[PSs:[PS:[Ps:1998,3,\epsilon], [Pf:1998,5,\epsilon]]]$  である。

### 3.2 概念の連結関数

以上で定義したように、フレームは互いに連結することで新しいフレームを構成する。4章で提案する解

析アルゴリズムでは、この連結の規則を利用して、実行されるので、本節ではその準備として、2つのフレームを引数とする連結関数を定義する。この定義では、補足概念のフレームも数多く登場する。このフレームは前節の最初で述べたとおり、その内部に図2の概念の略称を記述したものである。以後、特に断らない限り  $t, p$  はそれぞれ《時間》, 《時点》のフレームを表す。また、関数名は構成する概念名の略称の前に  $F$  を付けて表す。

#### (1) 《時間》を構成する関数

$$FT([T:y_1, m_1, d_1], [T:y_2, m_2, d_2])$$

$$= [T:y_1 + y_2, m_1 + m_2, d_1 + d_2]$$

例) 「1年」と「3ヵ月」のフレーム  $[T:1,0,0]$ ,  $[T:0,3,0]$  は、 $FT([T:1,0,0], [T:0,3,0])$  により連結され、「1年と3ヵ月」のフレーム  $[T:1,3,0]$  を構成する。

#### (2) 《時点》を構成する関数

$$FP([P:y_1, m_1, d_1], [P:y_2, m_2, d_2])$$

$$= \begin{cases} [P:y_1, m_2, d_2] & y_1 \neq \epsilon, m_1 = \epsilon, d_1 = \epsilon \\ [P:y_1, m_1, d_2] & y_1 \neq \epsilon, m_1 \neq \epsilon, d_1 = \epsilon \\ [P:y_1, m_1, d_1] & y_1 \neq \epsilon, m_1 \neq \epsilon, d_1 \neq \epsilon \end{cases}$$

$$FP([P:y_1, m_1, d_1], [T:y_2, m_2, d_2])$$

$$= [P:y_1 + y_2, m_1 + m_2, d_1 + d_2]$$

$$FP([Pb:p], [Tr:t]) = FP(p, t)$$

例) 「1998年から」と「2年後」のフレーム  $[Pb:[P:1998,\epsilon,\epsilon]]$ ,  $[Tr:[T:2,0,0]]$  は  $FP([Pb:[P:1998,\epsilon,\epsilon]], [Tr:[T:2,0,0]])$  により連結され、「1998年から2年後」のフレーム  $[P:2000,\epsilon,\epsilon]$  を構成する。

#### (3) 《時間区間》を構成する関数

$$FTS([Ts:ts], [Tf:tf]) = [TS:ts, tf]$$

$$FTS(t, [e:]) \quad (t_d \text{ は程度の幅})$$

$$= [TS:FT(t, -t_d), FT(t, t_d)]$$

例) 「約」と「10日間」のフレーム  $[e:]$ ,  $[T:0,0,10]$  は、程度の幅を  $t_d = [T:0,0,1]$  とするとき、 $FTS([e:], [T:0,0,10])$  により連結され、「約10日間」のフレーム  $[TS:[T:0,0,9], [T:0,0,11]]$  を構成する。

#### (4) 《時点区間》を構成する関数

$$FPS(p, [e:]) \quad (t_d \text{ は程度の幅})$$

$$= [PS:FP(p, -t_d), FP(p, t_d)]$$

$$FPS(p, [PS:ps, pf])$$

$$= [PS:FP(p, ps), FP(p, pf)]$$

$$FPS(p, [TS:ts, tf])$$

$$= [PS:FP(p, ts), FP(p, tf)]$$

$$FPS([Psb:p], t) = [PS:p, FP(p, t)]$$

$$FPS([Pfb:p], t) = [PS:FP(p, -t), p]$$

$$FPS([Pb:p], [TSr:ts]) = FPS(p, ts)$$

$$FPS([Ps:p_s], [Pf:p_f]) = [PS:p_s, p_f]$$

例) 「1998 年から」と「2000 年まで」のフレーム [Ps:[P:1998,ε,ε]], [Pf:[P:2000,ε,ε]] は  $FPS([Ps:[P:1998,ε,ε]], [Pf:[P:2000,ε,ε]])$  により連結され、「1998 年から 2000 年まで」のフレーム [PS:[P:1998,ε,ε],[P:2000,ε,ε]] を構成する。

(5) 《周期時点》を構成する関数

$$FPP([PP:p_b, t_p], p) = [PP:FPP(p_b, p), t_p]$$

例) 「毎年」と「4 月」のフレーム [PP:[P:1998,ε,ε], [T:1,0,0]], [P:ε,4,ε] は  $FPP([PP:[P:1998,ε,ε], [T:1,0,0]], [P:ε,4,ε])$  により連結され、「毎年 4 月」のフレーム [PP:[P:1998,4,ε],[T:1,0,0]] を構成する。

(6) 《周期時点区間》を構成する関数

《時点区間》を  $ps$  で表す。

$$\begin{aligned} FPPS([PP:p_b, t_p], ps) \\ = [PPS:FPPS(p_b, ps), t_p] \end{aligned}$$

例) 「毎年」と「春」のフレーム [PP:[P:1998,ε,ε],[T:1,0,0]], [PS:[P:ε,3,ε],[P:ε,5,ε]] は  $FPPS([PP:[P:1998,ε,ε],[T:1,0,0]], [PS:[P:ε,3,ε],[P:ε,5,ε]])$  により連結され、「毎年春」のフレーム [PPS:[PS:[P:1998,3,ε],[P:1998,5,ε],[T:1,0,0]] を構成する。

(7) 《時間》の下位概念を構成する関数

$$\begin{aligned} FTr(t, [rf:]) &= [Tr:t] \\ FTr(t, [rp:]) &= [Tr:-t] \\ FTs(t, [s:]) &= [Ts:t] \\ FTf(t, [f:]) &= [Tf:t] \\ FTp(t, [p:]) &= [Tp:t] \end{aligned}$$

例) 「1 年」と「前」のフレーム [T:1,0,0], [rp:] は  $FTr([T:1,0,0],[rp:])$  により連結され、「1 年前」のフレーム [Tr:[T:-1,0,0]] を構成する。

(8) 《時点》の下位概念を構成する関数

$$\begin{aligned} FPb(p, [b:]) &= [Pb:p] \\ FPS(p, [s:]) &= [Ps:p] \\ FPF(p, [f:]) &= [Pf:p] \\ FPSb(p, [sb:]) &= [Psb:p] \\ FPfb(p, [fb:]) &= [Pfb:p] \end{aligned}$$

例) 「1998 年」と「から」のフレーム [P:1998,ε,ε], [s:] は  $FPb([P:1998,ε,ε],[s:])$  により連結され、「1998 年から」のフレーム [Pb:[P:1998,ε,ε]] を構成する。

(9) 《時間区間》の下位概念を構成する関数

《時間区間》を  $ts$  で表す。

$$\begin{aligned} FTSr(ts, [rf:]) &= [TSr:ts] \\ FTSr(ts, [rp:]) &= [TSr:-ts] \\ FTSS(ts, [s:]) &= [TSS:ts] \\ FTSf(ts, [f:]) &= [TSf:ts] \end{aligned}$$

例) 程度の幅を  $t_d = [T:1,0,0]$  とするとき、「約 2 年」

と「後」のフレーム [TS:[T:1,0,0],[T:3,0,0]], [rf:] は、 $FTSr([TS:[T:1,0,0],[T:3,0,0]], [rf:])$  により連結され、「約 2 年後」のフレーム [TSr:[TS:[T:1,0,0],[T:3,0,0]] を構成する。

(10) 《時点区間》の下位概念を構成する関数

《時点区間》を  $ps$  で表す。

$$\begin{aligned} FPSb(ps, [b:]) &= [PSb:ps] \\ FPSs(ps, [s:]) &= [PSs:ps] \\ FPSf(ps, [f:]) &= [PSf:ps] \end{aligned}$$

例) 「1998 年春」と「から」のフレーム [PS:[P:1998,3,ε],[P:1998,5,ε]], [s:] は  $FPSs([PS:[P:1998,3,ε],[P:1998,5,ε]], [s:])$  により連結され、「1998 年春から」のフレーム [PSs:[PS:[P:1998,3,ε],[P:1998,5,ε]] を構成する。

## 4. 時間表現の解析アルゴリズム

### 4.1 形式表現への変換

本論文では、時間表現をそれに対応するフレームに変換することをその解釈と考える。実際の処理では、時間表現を入力成分列で表現し、これを長さが 1 のフレーム列に変換する。入力成分列は有限個の基本成分からなる列であり、その最後には末尾記号 \$ が付加される。また、フレーム列は有限個のフレームからなる列であり、それを構成するフレームの個数を長さと呼ぶ。特に、長さが 0 のフレーム列を空フレーム列と呼び、 $\lambda$  で示す。

変換処理の基本的な動作は、構文解析における Shift-Reduce 法<sup>1)</sup>と同じである。この特徴を文脈自由文法の用語を用いて説明すると次のとおりになる。

- (1) 非終端記号と前終端記号にはフレームが対応し、終端記号には基本成分が対応する。
- (2) 辞書規則は、基本成分をフレーム列に変換することに対応し、入力関数で与えられる。
- (3) 句構造規則は、構成関係にあるフレームを連結関数を用いて 1 つにまとめることに対応し、行動表で与えられる。
- (4) ユーザが設定するパラメータとして、参照時間と時間幅があり、入力関数と連結関数で使われる。
- (5) 処理過程において、フレーム列の格納にはスタックとディーキューが使用される。複数のフレーム列の候補が出現する場合、その数に応じてスタックとディーキューを用意し、それぞれの候補について順番に処理を進める。

### 4.2 入力関数

入力関数は、引数として基本成分をとり、戻り値と



reduce  $FPb$  より、「1998 年から」に対して同様に  $[Pb:[P:1998,\epsilon,\epsilon]]$  を得ることができる。しかし、このフレームと「2000 年」のフレーム  $[P:2000,\epsilon,\epsilon]$  に対しては、 $ACTION(Pb,P) = remove$  であるので、これらは接続不可能であることが分かる。このように成分が同じでも、対応する概念によって他の概念と接続できるものとできないものが存在する。

#### 4.4 参照時点と時間幅

参照時点と時間幅は、パラメータとしてユーザによって事前に与えられる。

参照時点は、時間表現が発話されたときに参照されていた時点であり、入力関数や連結関数の中で未定義の属性値があるときにそれを補完するために使用される。たとえば、「春」には、 $[PS:[P:\epsilon,3,\epsilon],[P:\epsilon,5,\epsilon]]$  のようなフレームが辞書に対応付けられているが、月属性より大きな単位の年属性が未定義であるので、参照時点の年属性の値をこのフレームの年属性の値に代入する。したがって、参照時点が  $[P:1998,4,20]$  の場合、「春」に対応するフレームは  $[PS:[P:1998,3,\epsilon],[P:1998,5,\epsilon]]$  になる。

時間幅は、程度の幅を表す時間であり、《時間》に対するフレームによって表す。このパラメータは、《程度》のフレームと他のフレームが接続するときに連結関数の中で使用されるが、その際は最小単位の属性値に対してだけパラメータの値を使用する。たとえば、「約」と「10 日間」のフレーム  $[e:]$  と  $[T:0,0,10]$  の連結において、時間幅を  $[T:1,1,1]$  としている場合、 $[T:0,0,10]$  の最小単位の属性は日なので、時間幅の日単位の属性値が程度の幅を表す時間として選択され、程度の幅として使用される《時間》のフレームは、 $[T:0,0,1]$  となる。

#### 4.5 変換アルゴリズム

入力成分列をそれに対応するフレームに変換する手続き TRANSLATE を次に示す。この手続きでは、処理中に生成されるフレーム列を一時的に格納する場所としてスタック集合 ( $S\_SET$ ) とディーキュー集合 ( $D\_SET$ ) を使用する。前者はスタックを要素とする集合であり、フレーム列はこのスタックに 1 つずつ格納される。後者もスタックがディーキューであるだけで前者と同様である。なお、ディーキューとは、その両端でデータの追加と取り出しが可能な列のことである。また、解析の初期状態 (initial state) と終了状態 (end state) を示すフレームとしてそれぞれ、 $[is:]$  と  $[es:]$  を導入する。これらと他のフレームと関係は表 1 の行動表に記述してある。また、これらに対する連結関数は、 $[is:]$  および  $[es:]$  と連結するフレームをそのま

ま戻り値として返すものが多いので、これらの定義は割愛する。特に、 $[es:]$  は  $INPUT(\$)$  から得られる。

#### 【手続き TRANSLATE】

【入力】: 入力成分列  $x = x_1, x_2, \dots, x_n (= \$)$

【出力】: 長さが 1 のフレーム列集合

【方法】:

(T1) 処理対象の基本成分の位置を  $i$  で表し、これを 1 に設定する。スタック集合  $S\_SET$  とディーキュー集合  $D\_SET$  をそれぞれ、 $\{[is:]\}$ ,  $\emptyset$  に初期化する。

(T2)  $i \leq n$  が成立する間、(T3) から (T7) までの処理を繰り返す。この後、(T8) に進む。

(T3)  $D\_SET$  に  $INPUT(x_i)$  の返すフレーム列の集合を格納する。 $D\_SET$  には、フレーム列と同数のディーキューが確保され、各フレーム列はそのディーキューの末尾に 1 つずつプッシュされる。

(T4)  $S\_SET$  の要素である任意のスタックを  $S$ ,  $D\_SET$  の要素である任意のディーキューを  $D$  とするとき、 $S\_SET \times D\_SET$  のすべての要素  $(S, D)$  に対して、(T5) および (T6) の処理を実行する。この後、(T7) に進む。

(T5)  $D \neq \lambda$  が成立する間、(T6) の処理を繰り返す。この後、(T4) に戻る。

(T6) スタック  $S$  の先頭のフレームとディーキュー  $D$  の先頭のフレームをそれぞれ、 $[c_s:v_s]$ ,  $[c_d:v_d]$  とする。行動表から、 $c_s$  と  $c_d$  に該当する項目を参照し、その種類に応じて次の行動を実行した後、(T5) に戻る。

(1) **reduce**  $f$  スタック  $S$  とディーキュー  $D$  の先頭から  $[c_s:v_s]$  と  $[c_d:v_d]$  をポップする。この後、 $S = \lambda$  ならば、連結関数  $f$  によって新しく生成したフレームをスタック  $S$  にプッシュする。そうでなければ、ディーキュー  $D$  の先頭にそれをプッシュする。

(2) **connect** ディーキュー  $D$  の先頭から  $[c_d:v_d]$  をポップし、これをスタック  $S$  にプッシュする。

(3) **remove** スタック  $S$  とディーキュー  $D$  を消去する。

(T7)  $i$  に 1 を加え、 $D\_SET$  を  $\emptyset$  にした後、(T2) へ戻る。

(T8) スタック集合の中から長さが 1 のフレーム列だけを出力し、この手続きを終了する。

【手続き終了】

入力成分列  $x =$  「1998 年」, 「から」, 「2000 年」, 「まで」, 「\$」に対する処理を説明する。この処理



表 2 動作例  
Table 2 An example of process.

ステップ	スタック集合 $S\_SET$	ディーキュー集合 $D\_SET$	処理対象の 基本成分	処理内容
1	[is:]	$\phi$	1998年	T3
2	[is:]	[P:1998]	1998年	T4
3	[is:](=S)	[P:1998](=D)	1998年	T6(reduce FP)
4	[P:1998]	$\lambda$	1998年	T7
5	[P:1999]	$\phi$	から	T3
6	[P:1999]	[s:] [b:]	から	T4
7	[P:1999](=S) [P:1999]	[s:](=D) [b:]	から	T6(reduce FPS)
8	[Ps:[P:1999]] [P:1999](=S)	$\lambda$ [b:](=D)	から	T6(reduce FPb)
9	[Ps:[P:1999]] [Pb:[P:1999]]	$\lambda$ $\lambda$	から	T7
10	[Ps:[P:1999]] [Pb:[P:1999]]	$\phi$	2000年	T3
11	[Ps:[P:1999]] [Pb:[P:1999]]	[P:2000]	2000年	T4
12	[Ps:[P:1999]](=S) [Pb:[P:1999]]	[P:2000](=D) [P:2000]	2000年	T6(connect)
13	[Ps:[P:1999]][P:2000] [Pb:[P:1999]](=S)	$\lambda$ [P:2000](=D)	2000年	T6(remove)
14	[Ps:[P:1999]][P:2000]	$\lambda$	2000年	T7
15	[Ps:[P:1999]][P:2000]	$\phi$	まで	T3
16	[Ps:[P:1999]][P:2000]	[f:]	まで	T4
17	[Ps:[P:1999]][P:2000](=S)	[f:](=D)	まで	T6(reduce FPf)
18	[Ps:[P:1999]](=S)	[Pf:[P:2000]](=D)	まで	T6(reduce FPS)
19	[PS:[P:1999]][P:2000]	$\lambda$	まで	T7
20	[PS:[P:1999]][P:2000]	$\phi$	\$	T3
21	[PS:[P:1999]][P:2000]	[es:]	\$	T4
22	[PS:[P:1999]][P:2000](=S)	[es:](=D)	\$	T6(reduce FPS)
23	[PS:[P:1999]][P:2000]	$\lambda$	\$	T7

スタックSは右側が先頭、ディーキューDは左側が先頭で右側が末尾

で使用する変数の値を表 2 に示す。スタック集合とディーキュー集合の項目では、点線で区切られた四角の部分はその要素であるスタックとディーキューを表している。また、各変数の値は、処理内容に記された手順を実行した後のものである。なお、以降では説明を簡単にするため、入力関数によって返される各成分に対するフレーム列集合をそれぞれ、{[P:1998,ε,ε]}, {[b:],[s:]}, {[P:2000,ε,ε]}, {[f:]}とし、さらに《時点》のフレームをそれぞれ、[P:1998], [P:2000]と略記する。また、FP([is:],[P:1998])とFPS([PS:[P:1998],[P:2000]], [es:])はそれぞれ、{[P:1998]}と{[PS:[P:1998],[P:2000]]}を返す。

まず、(T1)で変数の初期化を行う。(T2)では、 $i(=1) \leq 5$ なので、(T3)に進み、関数INPUT(1998年)が返すフレーム列の集合 {[P:1998]}をD\_SETに格納する。(T4)で求めたS\_SET × D\_SETの要素([is:],[P:1998])に対して、(T5)および(T6)を実行する。(T5)では、 $D \neq \lambda$ となるので、(T6)に進む。(T6)では、スタックSの先頭のフレーム [is:]とディーキューDの先頭のフレーム [P:1998]に対して行動表を参照する。該当する項目ACTION(is,P)はreduce FPである

るので、2つのフレームをポップし、FP([is:],[P:1998])より得られたフレーム列 [P:1998]をスタックSにプッシュし、(T5)に戻る。(T5)では、 $D = \lambda$ より、(T4)に戻る。(T4)では、S\_SET × D\_SETのすべての要素に対して(T5)および(T6)を実行したので、(T7)に進み、D\_SETを $\emptyset$ にする。この後、変数iに1を加えて、(T2)に戻り、(T2)の条件式が成立する間、同様の処理を継続する。そして、 $i = 6$ で、(T2)の条件式が不成立になるので、(T8)に進み、S\_SETの要素のうち長さが1のフレーム列 [PS:[P:1998],[P:2000]]を出力してこの手続きを終了する。

以上は手続きの流れの説明であるが、次に詳しい還元の様子を表2のステップ16から説明する。ステップ16の(T4)の処理では、S\_SET × D\_SETの要素は、([Ps:[P:1998]][P:2000],[f:])だけである。この要素に対して、(T5)および(T6)を実行する。(T5)では、 $S = [Ps:[P:1998]][P:2000]$ 、 $D = [f:]$ となり、 $D \neq \lambda$ であるので、(T6)に進む。(T6)では、スタックSの先頭のフレーム [P:2000]とディーキューDの先頭のフレーム [f:]に対して行動表を参照する。該当する項目ACTION(P,f)はreduce FPfであるの

で、2つのフレームをポップし、 $Fpf([P:2000],[f:])$ より得られたフレーム列  $[Pf:[P:2000]]$  をディーキュー  $D$  の先頭にプッシュする。その後、(T5)に戻るが、 $D \neq \lambda$  より、再び (T6) に進む。(T6) では、スタック  $S$  の先頭のフレーム  $[Ps:[P:1998]]$  とディーキュー  $D$  の先頭のフレーム  $[Pf:[P:2000]]$  に対して行動表を参照する。該当する項目  $ACTION(Ps,Pf)$  は  $reduce\ FPS$  であるので、2つのフレームをポップし、 $FPS([Ps:[P:1998]],[Pf:[P:2000]])$ より得られたフレーム列  $[PS:[P:1998],[P:2000]]$  をスタック  $S$  にプッシュする。その後、(T5)に戻り、 $D = \lambda$  なので、(T4)に戻る。

## 5. 評価

### 5.1 実験データ

提案手法では、形態素解析結果に対して、手続き  $TRANSLATE$  の入力となるフレーム列を関数  $INPUT$  で生成する。「1998年」の形態素解析では数字と接尾辞に分解されるが、関数  $INPUT$  では、この形態素列(成分)からフレーム  $[P:1998,0,0]$  を出力するので、形態素辞書の接尾語「年」には、年に関する値を抽出するための属性を付加しておく必要がある。さらに、「元日」や「大晦日」のように時間の数量情報が陽に表れていない語彙に対しては、語彙自身が持つフレーム情報を形態素列を成分に変換する辞書に追加しておく必要があるので、文献6), 16), 17)を用いて、この作業を行った。表3に概念とその個数、そして、その具体例を示す。表3の情報により関数  $INPUT$  は容易に得られる。表3では、時点と時点区間に対する語彙が多くなったが、これは特定の時点や特定の時点区間に多くの別称があることが原因である。

次に、解析対象となる時間表現の実験データを収集した。対象となるデータは、その差異をできるだけ少なくするために、4種類の分野から選んだ。収集の対象となったデータの種類の、新聞(新聞一面記事11日分)、雑誌(日経パソコン3冊分の記事の中からニュースや新製品などの報告記事)、問合せ文(文書データに対する13人の検索要求)、電子メール(研究室メンバーの交信文書)である。

### 5.2 実験結果

収集したデータに対して、提案手法の表現形式への変換実験を机上で行い、その結果を次の項目に分類して評価した。

- A: 出力が一意に正しい形式である。
- B: 複数の出力があるが、すべて正しい形式を含む。あるいは、一部正しい形式を

表3 登録単語の個数

Table 3 The number of words.

概念	個数	具体例
時間	53	半月, 両日, 旬月
時点	526	こどもの日, 大晦日, 明日
時間区間	66	小半時, 一両日, 旬余
時点区間	702	春, 第一四半期, 中旬
周期時点	73	毎日, 日曜日, 毎正時
周期時点区間	30	ウィークデイ, 毎週, 週明け
全体	1,450	

表4 実験結果

Table 4 The results of experiment.

文書	個数	判定			
		A	B	C	A+B
新聞	632	55.9%	37.2%	7.0%	93.0%
		535	253	44	558
雑誌	735	43.4%	51.6%	5.0%	95.0%
		319	379	37	689
問合せ	126	87.3%	5.6%	7.1%	92.9%
		110	7	9	117
電子メール	502	45.8%	45.2%	9.0%	91.0%
		230	227	45	457
全体	1,995	58.1%	34.9%	7.0%	93.0%
		1,012	848	135	1,860

含む。

C: 正しい形式が得られない。

すべての時間表現について変換処理を実施するのはかなりの労力を要するので、この実験では、時間表現を構成パターンに応じて数十種類のグループに分類し、各グループに対して代表となる時間表現を無作為に1つ選択した。そして、それらに対して変換処理を行い、その判定をグループ内全体の時間表現に対する判定とした。たとえば、「94年から」と「97年10月から」の場合、それら基本成分は《時点》と《開始》のように同じ構成パターンをしているので、同一のグループに含まれる。今回の実験におけるすべての処理は、2人で行い、3カ月の期間を要した。実験に際しては誤変換などを防ぐために、2人が出した結果を別の人が判定するという作業を行った。判定ごとの百分率とデータ数を文書別にまとめた結果を表4に示す。

いま、参照時点を1998年2月6日とするとき、それぞれの判定例を次に示す。

A: 「九三年三月から昨年五月にかけて」

$[PS:[P:1993,3,e],[P:1997,5,e]]$

B: 「2月6日」

$[P:1998,2,6],$

$[PP:[P:1998,2,6],[T:0,1,0]]$

C: 「あの日から三年」

なし

B判定では、「2月6日」について、《時点》と《周期時点》の2つの意味が解釈されるが、それぞれの解釈結果はどれも正しい。さらに、解釈を絞り込むには、文全体、あるいは文脈情報を利用する必要がある。また、B判定とされた結果は、先の2つの概念に《時間》が加わって出力されるものがほとんどであり、その出力数の分布は2個がB判定の中で53%、3個が47%であった。たとえば、「6日」の場合、先の2つの出力結果に《時間》としての[T:0,0,6]が加わる。不正解と判定されたものは、「11年」のような年号の値（この場合は平成）を表したものである。今回の実験では、西暦しか考慮していなかったためにそれ以外の年の値はすべて不正解とした。しかし、このようなものに対しては、年号に対する処理を加えることにより、正解へと改善することができると考えられるので、B判定と評価することにした。このような観点から、表4ではA判定が58.1%、B判定が34.9%、C判定が7.0%、AとB判定の合計は93%となり、時間表現を局所的に解決する手法としては、良い結果が得られた。

表4の文書別の結果に関して、問合せ文でのA判定の割合が90%近くになり他の文書に比べて高いことが分かる。これは、問合せ以外の文書では、省略表現が多いのに対して、問合せの文書ではこのような省略表現が非常に少ないことに起因している。省略の例では、現時点が「1998年3月3日」の場合、「3月3日」や単に「3日」と表現することであり、「3日」は、《時点》と《時間》など複数の意味解釈の可能性がある。以上より、自然言語インタフェースの代表的な応用分野である問合せ文の解釈に対して、本手法は有効であるといえる。

### 5.3 解析失敗の原因

正しい表現形式が得られない場合は、次の原因が考えられる。

#### (1) 時間表現自体が複数の意味を持つ場合

たとえば、「春秋」は、「春」と「秋」の2つの意味を持ち、両方とも時点区間を表す。また、「7月19日～28までの毎日」は「7月19日～28まで」と「毎日」の2つの意味を持ち、しかもそれぞれが《時点区間》と《周期時点》を表す。これらの意味は意味表現形式の集合として表すことができる。しかし、提案手法では、意味表現形式の集合に対して考慮していないので、このような時間表現を取り扱うことができない。

#### (2) 指示語を含む時間表現の場合

「その当時」のように指示語「その」が指す事柄に含まれる時間的な概念そのものを表すものや、「その前年」のように「その」が指す事柄に含まれる時間的な

概念を解明しなくては、表現形式の値が決まらないものなどがある。提案した表現形式には、指示的な概念をどのように表すかが考慮されていないので、この時間表現はすべてC判定となる。

#### (3) 時間表現に主観的な概念がある場合

「長年に渡って」中の「長年」のように個人の主観によって表現形式の値が異なるものがある。このような概念は、同じ人物であっても時と場面によってその意味が変化するので、意味表現形式で表すのは不可能であると考えられる。

## 5.4 考 察

本論文では、日本語時間表現について議論したが、基本的な考え方は他言語の表現についても同様に応用できる。

失敗例の解決には、より精度の高い解析技術が必要不可欠である。もし、指示代名詞の照応が解決できるならば、前節の(2)は解決できる問題である。同様にB判定の曖昧性の解決には時間表現以外の情報が必要となる。たとえば、省略のある時間表現がある場合、話題となっている時点や期間などの文脈情報を用いれば、省略による曖昧性を解消することができる。また、時間表現が接続する語句や修飾する語句が分かれば、時点と時間のどちらかというような曖昧性を解消することができる。

## 6. ま と め

本論文では、時間表現を意味解釈するために、意味解釈を時点、時点区間などの概念を分類し、これらの分類に対して、時間表現に対応する形式表現を定義した。また、形態素列からなる時間表現から形式表現に変換するアルゴリズムを提案し、4種類の文書データから抽出した約2,000の時間表現を対象とした実験により、提案手法の有効性を実証した。

本論文で報告した実験は、人手により行ったので、その評価にはかなり主観的なものが含まれていると考えられる。今後の研究課題としては、まず、客観的な評価を行うために、本手法を実際に計算機システムにおいて実装したうえで前述した実験を再度実施し、改めて本手法の評価を行うことである。さらに、自然言語インタフェースや視覚的な文書要約システムに応用して、より実際的な評価を行うことである。また、形式表現から多様な時間表現文や翻訳文を生成する研究も行う予定である。時間だけでなく、大きさ、重さ、長さ、速さなどの他の数量表現についても同様な形式表現が有用であるので、これらに対しても調査を進める予定である。

## 参考文献

- 1) 長尾真ほか：自然言語処理，岩波書店 (1996).
- 2) 伊藤元之，久保 晋，伊東幸宏：イメージ情報を利用した機械動作説明文の理解，電子情報通信学会論文誌，No.6, pp.931-945 (1995).
- 3) 松下光範，飯田敏幸：時間表現に基づく意図理解のための一考察，情報処理学会自然言語処理研究会資料，Vol.123, No.7, pp.49-56 (1998).
- 4) 田村直良：要約過程の形式化と実現について，人工知能学会誌，Vol.4, No.2, pp.196-206 (1989).
- 5) 田村俊哉，田村直良：文章の表現形式に基づいた要約文章の生成について，情報処理学会自然言語処理研究会資料，Vol.92, No.1, pp.1-46 (1992).
- 6) 池原悟ほか：日本語語彙体系，岩波書店 (1997).
- 7) 飯盛可織，佐川雄二，大西 昇：「の」を含む名詞句の日英翻訳に対する用例ベースアプローチ，情報処理学会自然言語処理研究会資料，Vol.112, No.8, pp.53-59 (1996).
- 8) 金田一春彦，林 大，柴田 武：日本語百科大事典，大修館書店 (1988).
- 9) 金田一春彦：日本語動詞のアスペクト，むぎ書房 (1976).
- 10) Allen, J.F.: Towards a General Theory of Action and Time, *Artificial Intelligence*, Vol.23, No.2, pp.123-154 (1984).
- 11) 原 裕貴，北上 始，中島 淳：時間概念の表現とデフォルト推論，人工知能学会誌，Vol.3, No.2, pp.216-223 (1987).
- 12) 玉野健一，松本祐治：制約条件を用いた事象の時間構造の記述，情報処理学会自然言語処理研究会資料，Vol.115, No.2, pp.9-14 (1996).
- 13) 西田元樹，松本祐治：テンス・アスペクトを考慮した現代日本語の複文生成，情報処理学会自然言語処理研究会資料，Vol.116, No.22, pp.151-156 (1996).
- 14) 上野晴樹，石塚 満：知識の表現と利用，オーム社 (1987).
- 15) 長尾 真：知識と推論，岩波書店 (1988).
- 16) 大野 晋，浜西正人：角川類語新辞典，角川書店 (1981).
- 17) 国立国語研究所：分類語彙表，秀英出版 (1964).

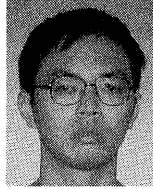
(平成 10 年 4 月 21 日受付)

(平成 11 年 7 月 1 日採録)



溝淵 昭二 (学生会員)

昭和 47 年生。平成 7 年徳島大学工学部知能情報工学科卒業。平成 9 年同大学院博士前期課程修了。現在同大学院博士後期課程在学中。情報検索，自然言語処理の研究に従事。



住友 徹

昭和 49 年生。平成 8 年徳島大学工学部知能情報工学科卒業。平成 11 年同大学院博士前期課程修了。現在同大学院博士後期課程在学中。自然言語処理の研究に従事。



泓田 正雄 (正会員)

昭和 46 年生。平成 5 年徳島大学工学部知能情報工学科卒業。平成 7 年同大学院博士前期課程修了。平成 10 年同大学院博士後期課程修了。現在同大学工学部知能情報工学科助手。情報検索，自然言語処理の研究に従事。



青江 順一 (正会員)

昭和 26 年生。昭和 49 年徳島大学工学部電子工学科卒業。昭和 51 年同大学院修士課程修了。同年同大学工学部情報工学科助手。現在同大学工学部知能情報工学科教授。この間コンパイラ生成系，パターンマッチングアルゴリズムの効率化の研究に従事。最近，自然言語処理，特に理解システムの開発に興味を持つ。著書「Computer Algorithms - Key Search Strategies」, 「Computer Algorithms - String Matching Strategies」IEEE CS press。平成 4 年度情報処理学会「Best Author 賞」受賞。工学博士。電子情報通信学会，人工知能学会，日本認知科学会，日本機械翻訳協会，IEEE，ACM，AAAI，ACL 各会員。