F-013

暦に基づく時間指示表現に対する意味表現形式
A Semantic Representation of Time Reference Expressions Based on Calendars

溝渕 昭二 †　　安藤 一秋 ‡
Shoji Mizobuchi　　Kazuaki Ando

1. はじめに

言語を媒体として人間とのコミュニケーションを図る情報処理システムにおいて，時間表現の意味を正確に理解することは非常に重要である．

時間表現には,「2001 年」や「今日」のように暦が規定する時間の区割りに基づいて特定の時間を指し示す表現が存在する．本稿では，これを「暦に基づく時間指示表現」，以降では，単に「暦表現」と呼ぶことにする．暦表現には，質と量の二つの意味的な側面がある．

暦表現を取り扱った従来の研究では，質的側面だけに着目した当該表現の認識方法 [1] や変換方法 [2] の提案がほとんどであり，量的側面までも含めた理解方法 [3] についてはあまり議論されていないのが現状である．

また，暦の時間モデルを対象とした研究には，本手法のベースとなった Leban らの手法 [4]，これより記述性の高い Ning らの手法 [5] などがある．しかしながら，そこで提案されたいずれの手法においても，限定された範囲の時間しか記述できない，あるいは，暦の区画にあった時間しか記述できないなどの問題点を持つ．

そこで，本稿では，暦表現を対象として，その意味の量的側面を記述するための手法を提案する．本手法では，暦表現の指し示す時間の範囲を時間指示式と呼ばれる式で表す．これは，従来，時点や期間，あるいは，それらの集合という形で別々に記述されてきた暦表現の示す時間の範囲を最終的にはコレクションという一つの数値情報に変換するものである．本手法により，暦表現の量的意味の記述と数値化が可能になる．

2. 暦表現

暦表現とは，暦に基づいて事態の歴史的な出現範囲を指し示す表現であり，時間表現の中では，最も具体的な内容を示す．暦表現の具体例を次に記す．

絶対暦表現	2001 年 3 月，8 月末日，日曜日，2003 年の第 7 週，1 月と 2 月
相対暦表現	今日，10 年後，2001 年から 2 年後，2 年ごと

暦表現には，絶対的なものと相対的なものの区分がある．前者は，時間を指示するときに基準とする時間が固定されているものであり，後者は，そうでないものである．本論文で対象とするのは，絶対暦表現とする．

暦表現の主要な役割は，事態の歴史的な出現範囲を指し示すことである．暦表現によって示されるその範囲が備える性質を考えた場合，少なくとも次の五つの性質が挙げられる．

1. 直線上の時間を指示すること
2. 幅を持つ時間を指示すること
3. 一つ以上の時間を指示すること
4. 暦をベースにして時間を指示すること
5. 統合的に時間を指示すること

なお，5 の「統合的」とは,「YYYY/MM/DD」等の記述形式や機能語によって時間を表す語句を組み合わせることを示す．

3. 暦表現の意味表現形式

暦表現の意味表現形式として時間指示式を提案する．時間指示式とは，時間軸上において定義されるコレクション，グラニュラリティー，演算子という三つの成分から構成される式である．式中の演算子が示す各種演算を適用することにより，時間指示式は最終的にコレクションに変換される．

以降では，時間軸，コレクション，グラニュラリティー，時間指示式について述べる．

3.1 時間軸

本手法では，数直線により時間を表すこととし，これを時間軸と呼ぶ．時間軸は，前章で述べた意味的な性質の 1 を満足するものである．

時間軸において，その単位を基本単位，時間軸の原点に対応付ける時間をゼロ時間と呼ぶ．例えば，基本単位を「日」とし，ゼロ時間を「2001 年 1 月 1 日」とすると，時間軸上の 1 に対応する時間は,「2001 年 1 月 2 日」となる．

以降では，特に断らない限り，基本単位とゼロ時間をそれぞれ,「日」と「2001 年 1 月 1 日」とする．

3.2 コレクション

コレクションは，暦表現が示す時間の範囲を表すものである．コレクションの定義の前にまずその要素となるインターバルを定義する．

時間軸上の相異なる二つの点を $l, u (l < u)$ とするとき，次のように定義される範囲 $\langle l, u\rangle$ をインターバルと呼ぶ．

$$\langle l, u\rangle = \{p \mid l \leq p < u\}$$

インターバルは，意味的な性質の 2 を満足するものである．なお，特別なインターバルとして，$\langle -\infty\rangle = \langle -\infty, -\infty\rangle$ と $\langle +\infty\rangle = \langle +\infty, +\infty\rangle$ を定義する

インターバルを要素とする集合において，相異なるインターバルの対がすべて重複しないとき，この集合をコレクションと呼ぶ．コレクションは，意味的な性質の 3 を満足するものである．

例えば,「2001 年」と「西暦」が指示する時間は，それぞれ $\{\langle 0, 365\rangle\}$ と $\{\langle -730485, +\infty\rangle\}$ で表される．

†近畿大学, Kinki University
‡香川大学, Kagawa University

コレクションについて，4種類の演算子（/[]，&，∼，|）を定義する．各演算子の説明と使用例を以下に記す．

/[] 演算子 $A/[n]B$ は，コレクション A の各インターバルに覆われるコレクション B のインターバルのうち，先頭 ($n<0$ のときは末尾) から n 番目のインターバルを要素とするコレクションを返す．

$$\{\langle 0,365\rangle\}/[1]\{\langle 0,31\rangle,\cdots,\langle 334,365\rangle\} \\ =\{\langle 0,31\rangle\}$$

& 演算子 $A\&B$ は，コレクション A の各インターバルに覆われるコレクション B のインターバルを要素とするコレクションを返す．

$$\{\langle 0,365\rangle\}\&\{\langle -366,-335\rangle,\langle 0,31\rangle,\langle 365,396\rangle\} \\ =\{\langle 0,31\rangle\}$$

∼ 演算子 $A\sim B$ は，コレクション A の各インターバルから，それと順序が同じ位置にあるコレクション B のインターバルまでを覆う最小のインターバルを要素とするコレクションを返す．

$$\{\langle 0,31\rangle,\langle 365,396\rangle\}\sim\{\langle 59,90\rangle,\langle 424,486\rangle\} \\ =\{\langle 0,90\rangle\ \langle 365,486\rangle\ \}$$

| 演算子 $A\mid B$ は，コレクション A と B のインターバルを要素とするコレクションを返す．

$$\{\langle 0,365\rangle\}\mid\{\langle 730,1095\rangle\} \\ =\{\langle 0,365\rangle,\langle 730,1095\rangle\}$$

3.3 グラニュラリティー

グラニュラリティーとは，時間を，周期的に連鎖し，かつ，整数によって識別可能なインターバルの系列に分割するものである．これは，意味的な性質の4を満足するものである．なお，インターバルに対応付ける整数をインデックスと呼ぶ．

周期の開始位置を p_s，p_s を含むインターバルのインデックスを i_s，周期内のインターバルの数を n，周期内の区割りを示すインターバルの幅の列を $w_1,\cdots,w_n$ とするとき，それらのパラメータによって定義されるグラニュラリティー G を次のように記す (図1参照).

$$G=\langle\langle p_s;i_s;w_1,w_2,\ldots,w_n\rangle\rangle$$

グラニュラリティーを G とするとき，インデックス i に対応付けられたインターバルを $G(i)$，インデックス $i_1,\cdots,i_n$ に対応付けられたインターバルを要素とするコレクションを $G\{i_1,\ldots,i_n\}$ で表すことにする．

週を表すグラニュラリティー W とそれを使って「毎週」を記述した例を次に示す（「*」はすべての整数を示す）．

$$W=\langle\langle 6;0;7\rangle\rangle$$

$$\begin{array}{ccccccc} \cdots & W(-2) & W(-1) & W(0) & W(1) & W(2) & \cdots \\ & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\ \cdots & \langle -8,-1\rangle & \langle -1,6\rangle & \langle 6,13\rangle & \langle 13,20\rangle & \langle 20,27\rangle & \cdots \end{array}$$

$$\begin{aligned} W\{*\} &= W\{\cdots,-1,0,1,\cdots\} \\ &= \{\cdots,\langle -1,6\rangle,\langle 6,13\rangle,\langle 13,20\rangle,\cdots\} \end{aligned}$$

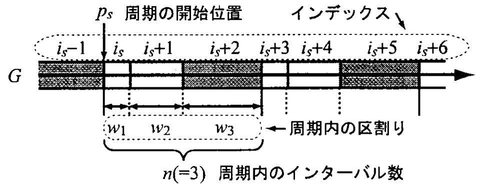

図1: グラニュラリティー

表1: 暦表現に対する時間指示式

西暦	$Y\{-2001\}\sim\{\langle+\infty\rangle\}(=AD)$
2001年3月	$AD/[2001]Y\{*\}\&Y\{*\}/[3]M\{*\}$
8月末日	$Y\{*\}/[8]M\{*\}\&M\{*\}/[-1]D\{*\}$
日曜日	$W\{*\}/[1]D\{*\}$
2003年の第7週	$AD/[2003]Y\{*\}/[7]W\{*\}$
1月と2月	$Y\{*\}/[1]M\{*\}\mid Y\{*\}/[2]M\{*\}$

3.4 時間指示式

時間指示式とは，グラニュラリティーとコレクション，そして，それらに対して定義された演算子から構成される記号列のことである．これは，意味的な性質の5を満足するものである．時間指示式は，演算を施すことにより，最終的にコレクションへと変換される．

いくつかの時間表現に対する時間指示式を表1に示す．この表に示したとおり，時間指示式を使えば，多様な暦表現の意味を記述することができる．

4. おわりに

本論文では，暦表現の意味表現形式として時間指示式を提案した．この式は，最終的にコレクションという数値情報に変換することができるので，量的な側面からも暦表現の意味を捉えることが可能になった．

今後の課題としては，まず，本手法の有効性を先行研究との比較等によって評価することである．また，本手法の適用範囲は，絶対暦表現に限られているので，「今日」や「2日後」などの相対暦表現についても拡張する必要がある．

参考文献

[1] 関根: テキストからの情報抽出, 情報処理学会会誌, Vol. 40, No. 4, pp. 370–373 (1999).

[2] 池原他: 日英機械翻訳のための時間表現の意味と対応関係の解析, 情報処理学会論文誌, Vol. 44, No. 2, pp. 451–465 (2003).

[3] 溝渕他: 日本語時間表現の一解釈法, 情報処理学会論文誌, Vol. 40, No. 9, pp. 3408–3419 (1999).

[4] Leban etc.: A Representation for Collections of Temporal Intervals (1986).

[5] Ning etc.: An Algebraic Representation of Calendars, *Annals of Mathematics and Artificial Intelligence*, Vol. 36, No. 1-2, pp. 5–38 (2002).