

自然言語動作概念のアニメーション表現について

倉田 志津 古川 忠昭 岡出 高徳 横田 将生
(福岡工業大学)

本研究は人間とコンピュータ間における自然言語と動画(アニメーション)によるコミュニケーションを可能とするインターフェース・システムを試作した。このシステムは心像意味論(MIDST:Mental Image Directed Semantic Theory)に基づいている。心像意味論とは自然言語、知識表現、心像(イメージ)、世界を包括的に取り扱う理論である。この理論に基づきシステムは、人間が入力する自然言語文を心像のモデルである抽象的な属性空間の軌跡として意味解釈し、環境条件等の制約を満足するように具体化した後、アニメーションに変換し出力するものである。本システムにより人体モデルを用いた動作のアニメーション化の実験を行った結果、かなり一般性のある処理が実現できる見通しを得た。

On Animation of Human Actions Conceptualized in Natural Language

Shizu KURATA Tadaaki FURUKAWA Takanori OKADE
Masao YOKOTA
Fukuoka Institute of Technology

The authors have constructed an experimental interface system which can realize human-computer communication using natural language and animation. The system is based on MIDST(Mental Image Directed Semantic Theory), which is such a total theory that deals with natural language, knowledge representation, mental image and worlds, including a methodology. According to this system, it builds up abstract (or conceptual) interpretations as loci in the abstract spaces, so called Attribute Spaces, it instantiates the interpretations in accordance with the environmental or contextual restrictions and, translates them into animations. Several animating experiments with human body models on this system have given a fairly good proof of generality of the processing.

1 はじめに

現在までに、自然言語の意味に関する研究が数多くなされてきている。しかしながら、自然言語理解処理における人間の理解過程を十分に説明でき、しかもコンピュータ処理にも適するようなものではなかった。そこで自然言語の意味を人間の

心像現象と対応づけ、コンピュータにも適した理論(心像意味論[1]と呼ぶ)の構築を行った。これは人間の自然言語解釈は心像(イメージ)に基づくものとし、心像のモデルを与えることにより自然言語意味記述および理解処理方法を提案しているものである。本研究では、コンピュータと人

間との自然言語と動画（アニメーション）によるコミュニケーションの実現を目的としている。そこで人間によって入力される自然言語文から、理解結果を視覚的なアニメーションとして表示することによるコミュニケーションをはかる。心像意味論により入力される自然言語文を心像のモデルである抽象的な属性空間の軌跡として意味解釈する。次にこの解釈結果の表現（軌跡式）を環境条件等の制約を満足するように具体化し、アニメーションに変換する。

本研究は人体モデルによる言語解釈のアニメーション化である。そのため、まず人体モデルに対しての動作に関する言語表現の一般化を行った。そこで以下では2章でここで行った動作に関する言語表現の一般化について説明し、3章でシステム構成、4章でシミュレーション、5章においてまとめについて述べていく。

2 動作に関する言語表現の一般化

人間の動作は後述する制約条件により図1のように区別することができる。ここでまず動作命令を以下のように一般化した。

(人間) が (人体の部位) を ([C]) 動かす。

[] は任意要素であることを示している。動作属性 (C) を表1に示す。但しこの表1において“V”はベクトル（多次元），“S”はスカラー（1次元）を表す。

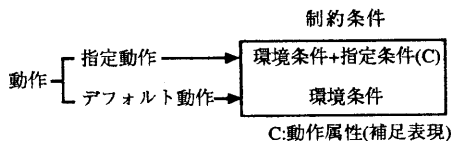


図1: 動作パターン

3 システム構成

システムは、自然言語理解モジュール、意味表現変換モジュール、アニメーション生成モジュールにより構成している(図2)。人間により入力さ

表1: 動作に関する属性

番号	属性	性質	属性表現例
A12	位置	[V]	p から q へ
A13	運動方向	[V]	上, 下
A14	向き	[V]	北, 南
A15	運動軌跡の形	[V]	ジグザク
A16	速さ	[S]	速い, 遅い
A17	移動距離	[S]	50cm
A38	回数	[S]	1回, 2回

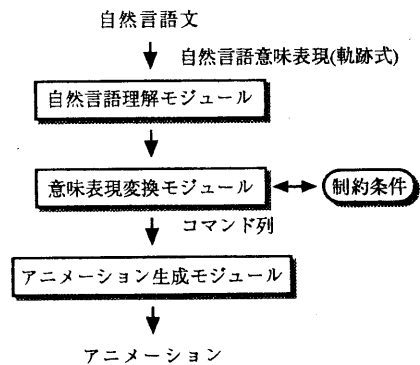


図2: システム構成

れる自然言語文を自然言語意味表現（軌跡式）に変換させ、次にその軌跡式を制約条件を考慮しコンピュータで処理できるコマンド列に具体化し、アニメーション化する。

3.1 自然言語理解モジュール

入力された自然言語を抽象的な属性空間の軌跡として意味解釈を行う。この解釈結果の表現（軌跡式）への変換を行う。

3.1.1 軌跡式変換

入力された自然言語文は初めに形態素解析を行う。これは最長一致法を用い、登録単語を参照して単語ごとに分ける。次に形態素解析の結果から構文解析を行う。ここでは語群形成規則を用い語

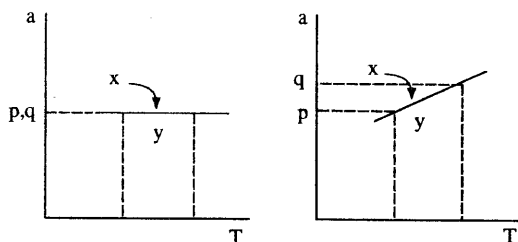


図 3: 原子軌跡式の模式図

群分割をする。最後に意味解析として軌跡式へ変換する [2],[3].

心像意味論の心像モデルにおいて、心像は属性空間(色立体や味覚立体に相当)と呼ばれる抽象空間における軌跡によって特徴づけられる。例えば、外界に実際に存在する事物の心像か否かは、「物理的場所」に関する属性空間に軌跡が存在するか否かで表示される。この属性空間に生じる時間的要因をもった軌跡のうち最も基本的なものとして、縦軸に属性空間(a), 横軸に客観的時刻(T)を設定することにより図3のように表現できる。このような軌跡は対象世界において時間または空間に関して単調(属性空間においては直線的)に継続または変化する事物の属性値およびその原因事物を表現するものである。なお、属性空間における時間とは客観的(物理的)時間ではなく心理的時間である。

図3で表現されるような軌跡を原子軌跡, それに対応する対象世界での事象を原子事象, および, それに対応する形式的システムでの式を原子軌跡式と呼び, 式(1)あるいはその簡約形(2)で与える。一般的な軌跡は, 原子軌跡で分節され原子軌跡式の論理的結合列で表現される。

$$L(x, y, p, q, a, g, k) \quad (1)$$

$$(x, y, p, q, a), a = (a, g, k) \quad (2)$$

ここで, L は述語定項, x および y は事物, p および q はそれぞれ時刻 t_i および t_j における事物 y の属性値 ($t_i < t_j$), a は属性, g は事象パターンの種類 ($g = Gt$ のとき時間的パターン, $g = Gs$

のとき空間的パターン), k は基準値, a は属性空間を表す。この軌跡式は, 軌跡式間の論理および時間的關係を同時に表示する時間的結合子を導入している。ここでは特に, 頻繁に用いられる同時的連言 (\square) および継時的連言 (\bullet) の定義を, それぞれ式(3), (4)で与えるに留める。

$$X_1 \square X_2 \triangle X_1 \wedge X_2 \wedge (t_1 = t_3 \wedge t_2 = t_4) \quad (3)$$

$$X_1 \bullet X_2 \triangle X_1 \wedge X_2 \wedge (t_2 = t_3) \quad (4)$$

但し, X_1, X_2 は任意の軌跡式であり, その時間帯は, 時間帯抽出関数 T を用いて次の式(5), (6)のように与えられているとする。

$$T(X_1) = [t_1, t_2] \quad (5)$$

$$T(X_2) = [t_3, t_4] \quad (6)$$

具体的な例を挙げて説明する。

(例) 「太郎は手をあげる」

この自然言語文に対する軌跡式は次のようになる。

locus(TARO migite ue ue a13)

この軌跡式の解釈は「‘TARO’の作用により‘migite’の属性がパラメータ3(‘ue’)からパラメータ4(‘ue’)への時間帯で, その対応する属性空間 a13 において単調な値域[‘ue’, ‘ue’]をとる。」となる。ここで自然言語文では‘手’となっているが軌跡式に変換した際に‘migite’としているのは, 一般的に‘手’という語彙を用いた場合‘右手’を指すことが多いと考えられるためである。また属性と値域への変換であるが, ‘あげる’という動作は属性番号‘a13’の属性空間において運動方向が上向き(‘ue’)となるためそれぞれに変換した。

例えばこの例の自然言語文に‘速く’という形容詞が含まれた「太郎は速く手をあげる」という自然言語文になると軌跡式は次のようになる。

locus(TARO migite ue ue a13)

\square locus(TARO migite hayai hayai a16)

3.2 意味表現変換モジュール

軌跡式を環境条件等の制約を考慮した上でコンピュータで扱いやすいコマンド列への変換を行う。

3.2.1 制約条件

一般に、人間の心(精神)の世界を内界、それが感知する物理的な世界を外界と呼んでいる。外界物は、物質や物体である物質体、および、それらが存在しうる物理的領域よりなる。外界事物は、そのような外界物が関与する事象で、物の移動や変形などを指示する。

外界事物である人間が、その存在する物理的領域で動作をする場合には図4で示したように、大きく分けて環境条件と指定条件の2つの条件があると考えられる。

環境条件

外界事物における物理的領域と、人体の部位の制約条件がある(図4)。部位の制約条件とは部位同士のつながりや、部位それぞれの可動性(動かせる方向、速さなど)である。

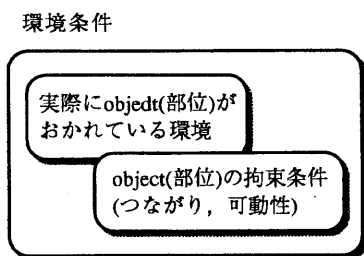


図4: 環境条件

(例) $connect(x,y) \wedge connect(y,z)$

$\wedge 手(x) \wedge 腕(y) \wedge 胴(z)$

指定条件

表1に示したような動作に関する属性(補足表現)を指す。

(例) 左へゆっくり歩く。

3.2.2 コマンド列

コマンド列を以下に示す。

```
{ repeat, return, number, frame,
  {{ part { xmax,xdou,xmin },
    { ymax,ydou,ymin },{ zmax,zdou,zmin }}}}
```

それぞれのパラメータは以下を表す。

- repeat : 繰り返し動作 (true:1, false:0)
- return : 往復動作 (true:1, false:0)
- number : 動作する部位数
- frame : 一動作のフレーム数
- part : 動作する部位番号
- xmax,xdou,xmin : 縦方向の最大角度, 縦方向1フレームの動作角度, 縦方向の最小角度
- ymax,ydou,ymin : 縦軸方向回転の最大角度, 縦軸方向回転1フレームの動作角度, 縦軸方向回転の最小角度
- zmax,zdou,zmin : 横方向の最大角度, 横方向1フレームの動作角度, 横方向の最小角度

3.2.3 コマンド列変換

本研究では人体モデルを障害のない空間に配置していると仮定した。そのため、環境条件としてはobject(部位)の拘束条件のみとなる。後、コマンド列への変換で考慮することは指定条件となる。これらの変換に際して具体的な例を挙げて説明する。

(例) locus(TARO migite ue ue a13)

この軌跡式に対するコマンド列は以下のようになる。

```
{0,0,1,18,{{3,{64,3.000000,0},
{0,0.000000,0},{0,0.000000,0}}}}
```

軌跡式は継時的連言でないため、repeat、return は 0 となる。また同時的連言でもなく、動作の対象となる事物は 'migite' のみであるため、number は 1 となる。さらに、速さに変化を持たせるような軌跡式もないため、frame はデフォルトで用意しているパラメータを代入している。動作は属性番号 'a13' の運動方向 ('ue') 上向きとなる。そこで現在の位置よりも縦方向において上への移動となるような動きをすればいいことになる。動作において上向きの動作という他のに、指定動作がないため、単純に縦方向のみの動きとしている。例えばこの例の軌跡式の他に、運動方向が 'migi' や 'hidari' などの語彙が含まれるようになると、部位の制約条件より縦軸方向回転の角度や横方向の角度などが加わりコマンド列が複雑になる。

本研究のシステムは、動きに幅を持たせるため動作範囲を可動可能範囲から乱数を用いて計算より求めている。そこで動作前の角度から動作させる方向へ角度を算出し、その 1 フレームの動作角度は frame を用いて求める。

以下に属性とそれに対応するコマンド列のパラメータを表 2 に示す。

表 2: 属性とパラメータの対応表

	位置	運動方向	向き	運動の軌跡	速さ	移動距離	回数
repeat	○			○			○
return	○			○			○
number		○	○		○		○
frame					○		
part	○	○	○	○			
xmax	○	○	○	○			
x0ou	○	○	○	○	○		
ymln	○	○	○	○			
ymax	○	○	○	○			
y0ou	○	○	○	○	○		
ymln	○	○	○	○			
zmax	○	○	○	○			
z0ou	○	○	○	○	○		
zmln	○	○	○	○			

3.3 アニメーション生成モジュール

表示ウィンドウは上部がアニメーション表示になっており、下部を 2 つに分け、一方をテキスト入力部、他方が軌跡式表示部になっている。アニメーション表示部はテンキーにより人体モデルを移動や回転できるようになっている。またテキスト入力部のテキスト処理には TextEdit という

Toolbox を採用している。

4 シミュレーション

「両手をゆっくりあげる」を入力した場合について説明する。

まず形態素解析を行う。

両手 | を | ゆっくり | あげる

次に構文解析を行う。

[両手, nc][を, jc][ゆっくり, Adj][あげる, vi] 語群に分類する。

名詞語群 | 副詞語群 | 動詞語群

これを軌跡式に変換するが、'両手' とあるのでこれを '右手' と '左手' の集合と等価変換する。またこの入力文では動作に関する言語表現を一般化した際の (人間) の指定がなされていない。このシステムではデフォルトとして 'イチロー' と名付けた人体モデルを指定するようにしているため、軌跡式における事物 x は 'イチロー' となる。さらに副詞語群を含む自然言語文のため、軌跡式は次のようになる。

locus(ICHIRO migite ue ue a13)

□locus(ICHIRO hidarite ue ue a13)

□locus(ICHIRO migite yukkuri yukkuri a16)

□locus(ICHIRO hidarite yukkuri yukkuri a16)

これをコマンド列に変換する。

{0,0,2,18

{3,{41,2.000000},{0,0.000000,0},

{0,0.000000,0}},

{5,{123,6.000000,0},{0,0.000000,0},

{0,0.000000,0}}}

右手と左手はそれぞれにパラメータを計算させているために、同じ最大角度になっていない。また、'a16' の 'yukkuri' という副詞が入っているため、デフォルトで用意されている一動作の角度よりも多く、つまりフレーム数を増やして動作をゆっくりさせている。以下にこのシミュレーションのアニメーション例を図 5 に示す。

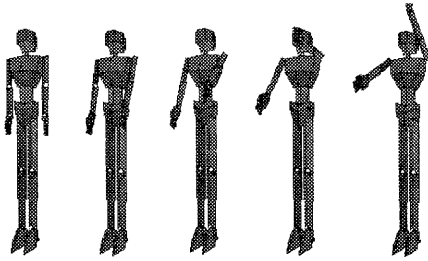


図 5: アニメーション例 1

この他に「タローは手を回す」と「頭を振る」について示す。

「タローは手を回す」を入力した場合であるが、‘回す’は属性番号‘a15’の運動軌跡が‘円’となる。そこで、軌跡式は以下ようになる。

locus(TARO migite en en a15)

このアニメーションを図 6 に示す。

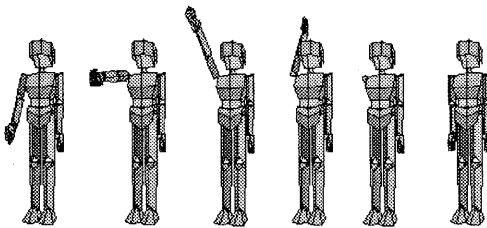


図 6: アニメーション例 2

また「頭を振る」を入力した場合であるが、‘左右’や‘上下’など属性番号‘a14’の‘向き’指定がなされていないためデフォルトとして‘左右’に振る動作とした。この場合の軌跡式を以下に示す。

locus(ICHIRO kubi migi migi a14)

● locus(ICHIRO kubi hidari hidari a14)...

‘振る’という動作は繰り返し動作と考えられるため、軌跡式は繰り返し出現可能であることを表す“...”となっている。また‘右(左)’に動作させ

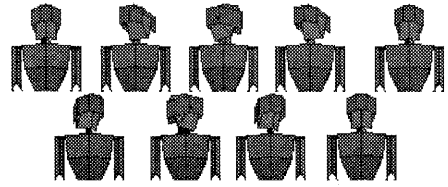


図 7: アニメーション例 3

た後に‘左(右)’に動作するので、軌跡式を継時的連言で結合している。

このアニメーションを図 7 に示す。

5 まとめ

人間の自然言語の解釈による心像(イメージ)は様々であるが、このシステムにより誰もが理解できる解釈結果のアニメーションによる出力ができた。本システムはまだ試作の段階で完璧な軌跡式への変換が行えるような単語登録や意味辞書を構築していない。しかし、同時に研究を進めている言語解析部分と接続することにより、もっと多様な動作のアニメーションが可能となる。これよりアニメーションに関する動作を増やし、多様な自然言語文に対応できるようになる。

本システムにより人体モダルを用いたアニメーション化の実験を行った結果、かなり一般性のある処理が実現できる見通しを得ることができた。これにより人間とコンピュータ間のコミュニケーションが可能になりつつあると言える。

参考文献

- [1] 横田将生:視覚情報と言語, 画像応用技術専門委員会サマーセミナー'95 Vol.4,1995.9.1~2
- [2] 横田将生, 心像意味論に基づくマルチメディア・コミュニケーション理論, 自然言語処理 106-10,1995
- [3] 横田将生, 西村靖司, 白石正人, 笠見一:心像意味論に基づく日本語連接名詞の構文および意味分析, 電子情報通信学会 (D),J77D,1,pp.131-142,1994