

TM-0563

自然言語入力による  
機器操作案内システム

浮田輝彦, 住田一男, 木下聡  
佐野洋, 天野真家(東芝)

June, 1988

©1988, ICOT

ICOT

Mita Kokusai Bldg. 21F  
4-28 Mita 1-Chome  
Minato-ku Tokyo 108 Japan

(03) 456-3191~5  
Telex ICOT J32964

---

**Institute for New Generation Computer Technology**

# 自然言語入力による機器操作案内システム\*

## An Operation Guidance System with Natural Language Input

浮田輝彦<sup>†</sup>

Teruhiko UKITA

住田一男<sup>†</sup>

Kazuo SUMITA

木下聡<sup>†</sup>

Satoshi KINOSHITA

佐野洋<sup>‡</sup>

Hiroshi SANO

天野真家<sup>†</sup>

Shin'ya AMANO

<sup>†</sup> 株式会社 総合研究所

Toshiba Corp. R&D Center

<sup>‡</sup> 新世代コンピュータ技術開発機構

Institute for New Generation Computer Technology

### 1. はじめに

ワードプロセッサを初めとして、オフィスにおける機器の機能は、ますます多様化、複雑化している。これらの機器を使いこなすためには、各種のコマンドに習熟する必要がある。そのためにマニュアルを読み、使用経験を積む必要がある。しかし、それでも暫く使わない期間があると、使い方を忘れてしまい、マニュアルを調べ直すことが良くおこる。マニュアルをオンライン化して、ディスプレイ上でマニュアルを調べる事が可能になってきている。しかし、適切なキーワードを指定するのが難しかったりメニューの選択が煩わしいなどの問題もあり、必要な情報を即座に取り出すことは難しい。本稿では、機器の使い方を忘れた場合などについて、自然言語を使って問合せできる機能について考察する。

自然言語処理技術を応用した質問応答システムとしては、積木の世界でロボットの腕に自然言語で動作を指示するSHRDLU<sup>(1)</sup>を初めとして、多くの実験的なシステムが開発されて来ている。コンサルテーションを対象とするものでは、オペレーティングシステムの一部機能の操作案内が知られている<sup>(2)</sup>が、入力文の処理が不十分である。例えば省略の表現が十分に扱えない、あるいは条件文が処理できないといった不十分

な点が残されている。自然言語文を理解するためには、対象とする機器の知識をシステムに持たせ、入力文中に現れる名詞や動詞に対応する概念を処理できる必要がある。

我々は自然言語で機器の操作法を問合せることができる質問応答システムの開発を目的として、フレームとルールの知識表現を基本とするコンサルテーションシステム (Information Service System by Analyzing Conversational Context - I S A C -) を開発している。本稿ではその概要を述べる。

### 2. 自然言語入力の特徴

一般の利用者が自然言語を入力として機械システムを使う場合の特長として、

- ①利用者が自分の言葉で入力できる。
- ②入力装置として安価なキーボードが使える。
- ③特別な訓練がいらない。

等が考えられる。反面、キーボードなどからの入力が難しい、という点があるが、オフィス等の機器を操作する際には、いずれにせよキーボードやそれに近いものを操作する必要がある。またワードプロセッサの普及により、今後キーボードからの自然言語入力は、ますます一般化していくものと考えられる。

このような自然言語を使って操作案内を行うときに必要と考えられるインタフェース機能には、次のものが考えられる。

- ①入力の会話文を省略なども含めて理解できること。

\*本研究は、ICOTからの再委託により、

第5世代コンピュータプロジェクトの一環として行っている。

②利用者の質問の意図を抽出し、適切に回答できること、

③システムが処理できない要求が入力されても、代替手段を提示すること、

更に利用者が初心者の場合には、次の機能が求められる。

④利用者にとって未知な操作の概念を教示できること、

⑤利用者の思い間違いによる入力に対する処置が可能なこと、

本稿では、これらの機能の中で自然言語入力の質問応答機能として最も重要な①の入力会話文の理解方法を中心に扱う。

### 3. システム構成

質問応答システムでは自然言語入力文の意味をシステムが持つ知識に照らして理解し、システムに課せられたタスクを実行する必要がある。図1にISACのシステム構成を示す。システムは知識処理を基本として、文解析・談話処理・タスク処理からなる。制御の流れは現在のところ単純なボトムアップになっている。また現行システムのタスク（応用）はビデオ装置を例として、その操作法に関するコンサルテーションとしている。

#### 3. 1 知識処理

入力されたメッセージが何を意味しているかを理解するためには、対象とするものの知識が必要であり、また推論の機能も要求される<sup>(3)</sup>。知識処理では、対象世界の知識を知識ベースとして表現し、他の処理部に推論機能を提供する。図2に知識表現の一部であるスキーマの例を示す。

す。本知識表現では深い知識と浅い知識を区別している。深い知識として、図2のようにビデオ装置の構成などはフレーム形式<sup>(4)</sup>で、また装置の動作や機能はフレームに付随したルール(srule, スキーマルールと呼ぶ)の形で表現している(これらをまとめてスキーマと呼ぶ)。フレームの部分には、その“概念”が他の概念に対して持っている関係をスロットとその値として記述してある。例えば、ビデオの上位概念(superC)は“電器製品”、部品として (has-part) “電源(power) ボタン、再生スイッチ”等を持つことを示している。スキーマルールでは、“ビデオの電源がオフで、power ボタンを

```
schema(cis.vir750x.  
  [(superC, 録画方式, [(value, [vhs ビデオ]))],  
    (superC, 録音方式, [(value, [hifiビデオ]))],  
    (has-part, [(value, [ powerボタン, 電源ランプ,  
                        再生スイッチ, ... ]))],  
    (ステイタス,  
      [(car-num,1),  
        (enumeration, [ 再生, 録画,  
                        待機, ... ] )]),  
    (電源, [(car-num,1),  
            (enumeration, [on,off]) ]),  
    [s-rule(dengen.dengen-on-1.V,  
      (kr-schema(V,[(電源,[off])],true),  
        event(押す,1, [(object,[SW])])),  
        kr-schema(V,[(電源,[on])],true),  
        [('$var-constraint',(  
          SW#power ボタン:-part-of(SW,V) ))]),  
      s-rule(dengen.dengen-lamp-1.V,  
        kr-schema(V,[(電源,[on])],true),  
        kr-schema(LP, [(点灯状態,[on])],true),  
        [('$var-constraint',(  
          LP#電源ランプ:-part-of(LP,V) ))])  
    ]),  
  ]),
```

図2 知識表現の例1 (ビデオのスキーマの一部)



図1 システム構成 (対象世界や経験的知識の表現, 推論機能)

```

rule( operation, play-back,
  seq(event(入れる2,1, [(object,[PW])]),
    event(入れる1,1, [(object,[K]),
      (goal,[VTR])]),
    event(押す,1, [(object,[SW])]),
    event(再生する,1, [(object,[K]),
      (tool,[VTR])]) ,
    [('$var-constraint',(
      VTR #vtr:
      K #カセット:
      PW#電源:-attribute -of(PW.VTR):
      SW#power ボタン:-part-of(SW.VTR) )]))
  ] ) .

```

図3 知識表現の例2 (浅いルールによる操作手順の表現)

```

schema(els, 録画する,
  [(superc, [(value, [動作])]),
  (agent, [(value-class, 人)]),
  (object, [(value-class, テレビ番組)]),
  (tool, [(value-class, vtr)]),
  (obligatory-case,
    [(value, [object, goal])]) ]),
[ ]),

```

図4 知識表現の例3 (イベントの表現)

押すという操作が行われると、電源がオンになる”という動作の例が示されている。

また図3に示すように、浅い知識として、コンサルティングのための経験的知識や機器の操作手順などをルールとして表現する。このレベルのルールはプロダクションシステム<sup>(5)</sup>として動作するようになっている。これによって、コンサルティングのための知識と対象世界の知識とが効率良く記述できる。

知識ベースには、さらに“出来事”として、「再生する」「動く」などの動作を示す概念（ほぼ動詞に対応している）についても、スキーマとして格フレームの形式で表現している。図4に例を示す。この例では、「録画する」の動作主には人が、対象にはテレビ番組がなりうること等が記されている。

知識処理部で利用できる推論機能は、ワールド機構を基本として、フレームにおける継承などの他にルールによる前向き・後向き推論が可能になっている。ルールによる推論では、スキーマのルール(srule)と因果関係のルールの2種類の推論が可能である。またワールド機能は、推論用のワーキングメモリをひとまとまりの

の(ワールド)として扱う機能で、仮説的な“世界”の中で推論を進めることや何か不具合が生じたときに元の状態に復帰できるメカニズムを提供する。この機能は、談話の処理や条件文を処理する際に用いる。

### 3.2 文解析<sup>(6)</sup>

文解析部は入力される漢字仮名混じりの文を形態素解析および構文解析し、入力の述語に対応して、格フレーム表現を出力する。まず入力文字列を文節に分割した後に、全文節間の係り受けの可能性を調べ、矛盾のないものを出力する。このとき係り受け関係の非交差と順方向性を制約条件として用いている。更に入力される発話に複数の述語が含まれるときには、それらの間を「条件・目的・時間前後関係・並列・因果」などの関係を記述した形式を出力する。

### 3.3 談話処理

会話文に対する談話処理として、特に重要である省略情報の補足と代名詞の同定を行い、入力発話の解釈を求める。省略された情報や代名詞の指示対象の判定では、知識ベースとの照合により、候補を選択する。

まず、指示名詞、“その”、“この”のついたもの、必須格の省略、名詞単独を照応型として検出する。これらの照応型について、過去の会話で出現した“もの”に対応するインスタンスを候補として抽出する。このとき知識ベース中の上位/下位の関係などを参照して候補を抽出する。例えば、「押す」の対象となるものは“ボタン”と知識ベースに記述されているときには、「それを押す」の“それ”は、“ボタン”に属する物のみ(ボタンの下位概念である再生ボタン、録画ボタンなど)を候補として取り出す。

次に、知識ベースに記述されている操作手続きなどのシステムのタスク処理に関する知識を参照して最適な候補を選択する。ここでは、知識ベースに記述されているはずの対象世界の動作間の因果的な関係(ビデオの場合、操作手続きに相当する)を調べることにより、最適な候補を判定する。例えば、「カセットをビデオに入れて再生ボタンを押したが( $\alpha$ が)動かない」という入力がある場合、 $\alpha$ の候補として、この

入力発話だけについてもカセット、ビデオ、再生ボタンのインスタンスが考えられる。入力されたもので最近傍の候補を選ぶと再生ボタンになるが、この入力の場合は、“ビデオが動かない（動作しない）”と解釈する方が好ましい。この判定には、対象の知識として記述される動作間の関係、すなわち今の例については、

もし “ビデオの電源を入れる”，かつ  
 “カセットをビデオに入れる”，かつ  
 “再生ボタンを押す”，  
 ならば “ビデオが動く”，  
 という操作手順のルールがあるので、これを参

```
[concept/入れる,
syntax/category/ 動詞型述語, source-word/入れた],
surface-case/
  に/[concept/ビデオ,
    syntax/category/ 準用詞型述語,
    source-word/ビデオに]],
を/[concept/カセット,
  syntax/category/ 補語,
  source-word/カセットを]],
fixed-portion/ [能動動詞, 肯定, 基本態, 過去, ... ]]
```

(a) 文解析の結果

```
[top/al,
al/[ [入れる, {goal/vtr1000,
  object/ カセット1001
  fixed-portion/ [肯定, 過去, ... ]}] ]]
```

(b) 談話処理の結果  
 (vtr1000 やカセット1001はインスタンスを示す)

図5 文解析・談話処理の出力例  
 入力文：「ビデオにカセットを入れた」

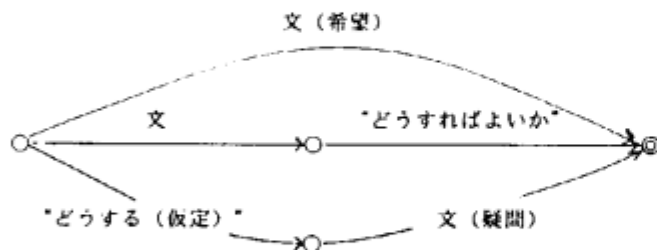


図6 発話の解析用オートマトンの例 (一部)

考にして、“α=ビデオ”と判定する。またこの時ある解釈の候補と適合するルールが複数存在する場合には、結論部が否定されるものを優先している(7)。

現在は先行詞が明示されている場合を取扱っている。談話処理の結果、必要な情報が補われた格フレーム表現が出力される。図5に文解析と談話処理の処理結果の例を示す。

### 3. 4 タスク処理

タスク処理は、一般にはシステムの応用（業務）処理を行う部分であり、ここでは操作法のコンサルテーションを行う。取り扱う質問のタイプは大きく分けて、

①装置の状態に関する単純な問い合わせ。

(例：「ビデオの電源はオフか」)

②目的-手段についての問い合わせ。

(例：「早送り再生するにはどうすればよいか」)

(「再生スイッチを押すとどうなるか」)

③操作ミスに起因する不具合の解決要求。

(例：「イジェクトボタンを押したが出ない」)

の3種類である。

タスク処理では入力された発話の文章をオートマトンによって解析して、その文章の構造に従って各々上記の3タイプの問題解決モジュールを起動する。図6に②のタイプのオートマトンの例を示す。例えば、入力が「どうすれば、ビデオを再生できるのか」のときは、オートマトンの下のルートを通る形で、受理される。そして「再生する」手順を、浅いルールの形式で記述されている知識ベースから検索し、「ビデオの電源ボタンを押して、カセットテープを入れ、再生ボタンを押してください」という応答文を生成する。

タイプ①の場合は、問合わせている対象となる概念を取り出し、その状態（スロットの値として記入されているもの）を取り出す。またタイプ③の不具合の場合は、そのときの希望する動作の手順をチェックして未だ行なわれていない動作を探し出して、利用者に確認する。このタイプの場合、操作順序のミスまでを扱うことにしている。

一方、これらのタイプとして認められなかつ

たものは、ほとんどのものが、利用者の行った動作の表現（例えば、「録画ボタンを押した」など）と考えられるので、それを知識ベースに追加して、スキーマルールと手順の浅いルールによる前向き推論を行うことにより、対象の機器の状態をシミュレートしておく。

入力された発話が、②のタイプで「・・・するとどうなるか」といった動作・結果の因果関係に関する質問の場合は、新しいワールドを構成し、その上で、前向き推論を起動する。その結果得られる状態の変化を応答として出力する。最後に、推論結果のタイプ別に用意される格フレームの表現から日本語文を生成している。

会話の制御方法として、スタックを使って基本的な入れ子構造に対処する。本システムでは利用者の入力は、どの時点においても自由な形式にしているので、システム側が利用者に質問をした場合でも必ずしもそれに応える必要はなく、他の問合せを発してもよい。処理の流れとして、利用者からの文が問合せであるときにはその要求をスタックにプッシュしてから、処理を進める。そして現在の間合せについての処理が終了すると、スタックをポップアップし、元の間合せ関する処理を再開する。

### 3. 5 動作例

本システムは逐次推論マシン P S I 上で試作している。現在知識ベースには、家庭用 V T R について、スキーマとして約 3 0 0 個の概念と 1 8 0 種類のスキーマルール、及び手続きなどを表現する浅いルールとして約 3 0 個が用意されている。これらは、V T R の操作マニュアルから抽出した。

図 7 に会話の例を示す。例 1 の会話では、単純な操作の問合せと操作の不具合の処理の例が示されている。また、例 2 では、条件文の処理の例が示されている。この場合、入力される明示的な時間関係の記述（その後、その前、代りに、など）に従って、ワールドを制御している。

## 4. おわりに

対象の知識を持ち自然言語入力による機器操作案内機能を持つ質問応答システム I S A C の概要を述べた。本システムの特徴として、

①省略表現を含む自然言語入力により機器の

ユーザ：ビデオでカセットを再生したい

システム：電源を入れて、カセットをビデオに入れて、再生スイッチを押して下さい。

ユーザ：カセットを入れて再生スイッチを押したが動かない

システム：ビデオの再生ランプが点灯していますか。

ユーザ：はい

システム：V H F 出力端子がテレビにつながっていますか。

ユーザ：いいえ

システム：V H F 出力端子がテレビにつながっていないからです。

#### (a) 会話例 1

ユーザ：カセットをビデオに入れた

システム：はい、わかりました。

ユーザ：再生ボタンを押したらどうなるか

システム：ビデオランプが点灯して再生が始まります。

ユーザ：その後で早送りボタンを押すと

システム：早送り再生が始まります。

#### (b) 会話例 2

図 7 試作システムによる会話例

（ビデオ、テレビがそれぞれ 1 台ずつあり、電源などはオフの初期状態を想定している）

操作案内を行う。

②ワールド機能を持ち、それにより条件文を処理する。

③知識ベースが手続きの記述と概念の体系記述に分けて表現できるので、対象機器の操作マニュアルなどから、比較的容易に知識ベースを構築できる。

がある。

新たな機器に対して本システムを適応させるためには、日本語解析の辞書と知識処理の知識ベースを変更する必要がある。現システムでは、ビデオの操作マニュアルだけから、必要な概念

と各種のルールを人手により抽出している。現在、知識ベースの一般的な標準化された表現手順は、確立されていないが、本システムでは、機器の構造と機能の側面に分けて記述できるようになっており、一般のエキスパートシステムと比較すると、記述に必要な労力は少ないといえる。

なお、操作案内の親切さの観点から見ると、案内応答すべてを、自然言語の文字列で表記するのは分りづらい場合がある。マニュアルについても、図やイラストによる表現が多くみられる。このような情報を説明するには、操作のポイントとなる点は自然言語で呈示し、必要に応じてイメージによる操作案内を併用することが現実的なアプローチと考えられる。今後、試作システムによる実験を通じ、改良して行きたい。

#### 参考文献

- (1) T.Winograd: "Understanding Natural Language", NY.:Academic Press (1972).
- (2) R.Wilensky, Y.Arens, and D.Chin:"Talking to UNIX in English:An Overview of UC", Comm. ACM, Vol.27, No.6,pp.574-593(1984).
- (3) 木下・佐野・浮田・住田・天野: "文脈理解のための知識の表現と推論", Proc. The Logic Programming Conf.'88, pp.205-212(1988).
- (4) M.Minsky:"A Framework for Representing Knowledge", In P.Winston(Ed.), The Psychology of Computer Vision, NY.:McGraw-Hill, pp.211-277(1975).
- (5) A.Newell and H.A.Simon:"Human Problem Solving", NJ.: Prentice-Hall (1972).
- (6) 佐野・杉村・赤坂・久保: "語構成に基づく形態素解析", 情報処理学会 自然言語研究会資料 NL 66-3 (1988.5).
- (7) 住田・浮田・木下・佐野・天野: "最適解判定法による会話文の理解", 信学技法NL C 87-16(1987.12).