

# インタラクティブ・システムにおける状況的な行為とその協調計算

小 窪 浩 明<sup>†\*</sup> 匂 坂 芳 典<sup>†</sup>  
 鈴 木 紀 子<sup>††</sup> 岡 田 美 智 男<sup>††</sup>

日常での会話やある秩序だった行為は必ずしも事前に用意されたプランや手続きによってナビゲートされているわけではない。最近のエスノメソドロジー研究なども指摘するように、会話などの相互行為は、状況的な行為とその文脈との相互作用の結果としてオリジナルに組織化され、レギュレートされているといえる。こうした視点から会話現象をとらえ直す作業と並行して、インタラクティブ・システムもまた「複雑で多様な相互行為を組織化する会話参加者の1人である」との立場から、その設計に対する見直しが必要に思われる。本論文では、状況的な行為とその文脈との相互作用により目的的な行為を組織化するメカニズムに基づいて、インタラクティブ・システムのプロトタイプを構築することを試みた。また、その基本的な振舞いを検証した結果、ここで構築されたプロトタイプシステムでは、(a)省略をともなう発話の解釈を文脈との協調計算やユーザとの相互行為の結果として組織化する、また、(b)文脈からの影響や状況的な行為（ここでは「ふるまい」と呼ぶ）の間での相互依存の度合いを制御することにより、それぞれ性質の異なる解釈をオリジナルに組織化する、などの柔軟な性質を備えることを明らかにした。

## Situated Actions and Cooperative Computation for Interactive Systems

HIROAKI KOKUBO,<sup>†\*</sup> YOSHINORI SAGISAKA,<sup>†</sup> NORIKO SUZUKI<sup>††</sup>  
and MICHIO OKADA<sup>††</sup>

Everyday conversation or systematic behaviors are not always navigated by predetermined plans and procedures. We consider that mutual behaviors such as conversation are sequentially organized and regulated as a result of interactions among situated behaviors, which we take the ethnomethodological literature as supporting this position. In parallel with the study of conversational phenomena from this viewpoint, we believe that it is necessary for an interactive system to be redesigned according to the stance that a designer of the system is a participant in a conversation and must organize various complex behaviors interactively. In this paper, we constructed a prototype interactive system that has a mechanism which meets these aims by organizing its behaviors as a function of the interaction between situated behaviors and the context. We also evaluated the behavior of the prototype system. We confirmed that the prototype system had the flexibility to be (a) organized into interpretations for utterance with omission as a result of interaction among users or cooperative calculation respect to the context, and (b) organized into different interpretations by controlling a parameter for mutual dependence between the context and situated behaviors.

### 1. はじめに

日常的な会話はさまざまな要因が関与し、複雑で多様な挙動を示す。そこで観察された会話の現象を分析

したり解析したりしようとする、その挙動の多様さに圧倒される。これまでの対話システムやインタラクティブ・システムの研究は、こうした多様な挙動をどのようにハンドリングするか、という視点で進められてきた。参照フレームの問題<sup>1)</sup>としてとらえれば、システム的设计者の多くは、すでに出現してしまった、あるいは出現しつつある多様な挙動を傍観者として観察し、そうした複雑な挙動をハンドリングするために、複雑な内部構造を備えたシステムを設計しようとしてきた。あるいは、そうした挙動をストキャスティックな過程としてとらえることを試みてきたといえる。し

<sup>†</sup> ATR 音声翻訳通信研究所  
ATR Interpreting Telecommunications Research Laboratories

<sup>\*</sup> 現在、日立製作所中央研究所  
Presently with Hitachi Central Research Laboratory

<sup>††</sup> ATR 知能映像通信研究所  
ATR Media Integration & Communications Research Laboratories

かしながら、「サイモンの蟻」の議論<sup>2)</sup>にもあるように、こうした挙動の複雑さ、多様さは系内部の複雑さの直接的な現れとは限らない。砂浜に残る蟻の足跡のように、単純な、状況に合わせた振舞いの相互作用によって、一見複雑に見える挙動は容易に作り出すことができる。その意味で、日常的な会話における複雑で多様な相互行為の結果を、会話参加者の複雑な内部構造に一方向的に帰属させることには問題がある。

一方、最近のエスノメソドロジー研究<sup>3)</sup>の指摘していることの1つは、日常的な会話における相互行為の手順や方略は外部からあらかじめ与えられたものではなく、むしろ「会話参加者によって相互行為的に、状況的にオリジナルに組織化されるものである」という点であろう。日常での何気ない行為や何気ない会話は、必ずしも事前に用意されたプランや手続きによってナビゲートされているわけではなく、むしろその文脈における状況的な行為が相互作用した結果として、オリジナルに組織化され、レギュレートされているといえる。こうした議論は状況論的な認知やロボティクスの分野などにおいて、1980年代の後半からなされてきたことでもある。

こうした視点は、ユーザとインタラクティブ・システムとの関係にも展開できる。インタラクティブ・システムもまた「複雑で多様な会話の挙動を組織化する会話参加者の1人である」という視点からその設計を見直していくことで、複雑で多様な会話の挙動に対するハンドリングの方略を変えていくことが期待できる。つまり、すでに出現してしまった挙動を傍観者として観察するのではなく、会話参加者の1人として系にかかわり、一緒に複雑な挙動を作り出すという立場への転換である。ここでは必ずしも複雑さや多様性の発現を一方向的にシステム側の内部構造に帰属させる必要はない。また、これまでのシステムの設計者はその系にとっての傍観者であって、必ずしも参加者ではなかったといえる。この設計者の参照フレームを系の内部に置くことはできないだろうか。こうした姿勢は、Normanら<sup>4)</sup>のいう「ユーザセンタード・デザイン (User-Centered Design)」においても指摘されてきたことである。

本研究は、インタラクティブ・システムもまたユーザとのかかわりの中で「複雑で多様な会話の挙動を組織化する会話参加者の1人である」との視点から、インタラクティブ・システムの基本的な枠組みに対する見直しの一環として行われているものである。特に、日常的な会話は文脈や環境に対する状況的な行為によって組織化されるという立場から、インタラクティブ・

システムとユーザとの間での相互行為としての問題解決行動を、各々の状況的な行為の相互作用によって出現させることを試みている。本論文では、インタラクティブ・システムとユーザとのかかわりにおける状況的な行為の出現を「ふるまい (behavior)」とよぶ複数の状況に埋め込まれたエージェント (situated agent) の相互作用によって実現することを試みる (本論文では、この状況的な行為 (behavior) をカッコ付きひらがなの「ふるまい」で表記し、一般用語である「振舞い」と区別して用いる)。また、駅の窓口などで交される簡単な会話を想定したプロトタイプシステムを構築し、いくつかのインタラクションの過程を例として、その基本的な振舞いとその性質について明らかにする。

以下、2章では、日常的な会話の相互行為としての性質についてまとめる。3章では、相互行為の中で状況的な行為を組織化する動的行動選択モデルの概要を述べる。4章では、プロトタイプとして構築された音声対話システムの概要とその動作原理について述べ、5章においてその振舞いの有効性について検証する。6章はまとめである。

## 2. 会話の組織化における状況的な行為

我々の日常での会話では、司会者などを特に必要とすることなく、発話の交替などが無意識に行われ、次々に話題が生じ展開していく。これらの振舞いや制御方略そのものは、さまざまな要素が状況的に振る舞うなかで、結果として立ち現れたものであり、制御が系の外から与えられたものではない<sup>5)</sup>。日常的な行為や日常的な会話などの相互行為は状況の中でオリジナルに組み立てられ、レギュレートされると考えられる。

ここで文脈のなかで状況的に振る舞うとはどういうことだろうか。日常での何気ない振舞いや何気ない会話を考えるとき、重要な視点の1つは「その何気ない行為は単独では意味を確定できない」という性質を備えることである。我々が散歩の途中で何気なく一步を踏み出すとき、その行為の意味や価値を考えながら歩くことはないだろう。その意味や価値は、その行為が大地とかかわる中で事後的に明らかにされる。我々はその行為の意味や価値をいったんは環境 (この場合には「大地」) に預ける、委ねることをするようである。その結果として、その行為は環境にナビゲートされる。この「自分の行為の意味や価値を環境に委ねながら、その環境にナビゲートしてもらおう」という基本的な姿勢は、生態心理学的なアプローチにおける行為に対するとらえ方や視覚的制御の考え方<sup>6)</sup>と重なる。また、

社会的な相互行為の場面で考えるならば、最近のエスノメソドロジー研究や対話研究において着目されつつある「ある発話とそのグラウンディング (grounding) の過程」においても共通した関係が見いだされる<sup>7)</sup>。

我々の日常的な会話において、単独の発話だけではその意味や機能は確定されない。その発話  $U_i$  の意味や価値をグラウンドするものは、その発話  $U_i$  に向けられた次の発話  $U_{i+1}$  である。こうした視点で考えるならば、最初の発話  $U_i$  は、他者の発話  $U_{i+1}$  によるグラウンディングを期待しつつ、自分の発話  $U_i$  に対する意味や価値を他者の応答  $U_{i+1}$  に預ける、委ねるといった基本的な姿勢を備えているといえる。また、その意味で発話  $U_i$  は自分の行為の意味や価値を知覚するための知覚行為を構成しているといえる。自分の行為の意味を知覚するためには、発話するという行為が不可欠であり、その行為を遂行するためには、その行為の可能性をつねに知覚する必要がある。

社会的な相互行為における状況的な行為においても、知覚と行為とが不可分の関係にあり、状況的な行為の循環的な連鎖が会話を組織化しているといえる。文脈 (たとえば相手の発話) だけに合わせたような、状況のみに依存させた主体性のない発話は、その会話の中でほとんど意味をなさないように思われる。しかし、こうした発話はむしろ日常での何気ない会話の雰囲気を作り出す<sup>8)</sup>。一方、課題指向的な会話や協調的な問題解決過程としてとらえられる会話も、話し手の抱くプランなどによって一方的に進められているわけではない。むしろ状況に合わせた発話をつなぎながら、話し手の目的を実現するべく有利な文脈を作り出していく。こうした過程を経ながら、結果として問題解決行動や合目的な行為を組織化していくものと考えられる。このことはインタラクティブ・システムの設計においても留意されるべきことであろう。

### 3. インタラクティブ・システムにおける行為の組織化

#### 3.1 動的行動選択モデル

日常的な会話などの相互行為に限らず、他者の発話を解釈する、ある発話を生成するなどの日常的な行為もまた状況的な行為と文脈との相互作用の結果として組織化されると考えられる。これらの振舞いを実現するために、本研究では「状況に埋め込まれたエージェント (situated agents)」の1つの実現形態である Agent Network Architecture (ANA) に基づいた動的行動選択モデル<sup>9)</sup>をインタラクティブ・システムに応用することを試みる (図1)。各エージェント

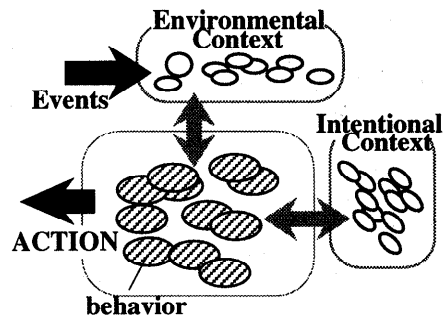


図1 インタラクティブ・システム概念図  
Fig.1 Concept of interactive system.

(situated agent) は、文脈の中で自己の「行為の可能性」を知覚し、行為を遂行する基本的な単位である。ここでは「ふるまい (behavior)」と呼ぶ。これは文脈に見い出される行為の可能性に従うだけでなく、文脈に影響を与える形で自律的に動作する。たとえば、自己の局所的なゴールを達成するために周囲の「ふるまい」と協調・競合の中から自分の行為に有利な文脈へと変えようと働きかけをする。ただ実際に自分の行為に有利なように文脈を変えられるか否かは、そうした試みの結果でしか知りえない。こうした基本的な振舞いは、「自分の行為の可能性を環境にいったんは委ねている」という解釈ができる。つまり、自分の行為の可能性や自分の役割をいったんは文脈に委ねながら、その結果として文脈によってナビゲートしてもらおう。そうした基本的な方略を備えている。

#### 3.2 2つの文脈

ここでは文脈を表現するため、「環境的な文脈 (environmental context)」と「意図的な文脈 (intentional context)」と呼ぶ2つの文脈をシステムの内部に用意した<sup>8)</sup>。いずれの文脈も動的な行動選択におけるリソースとして参照される。環境的な文脈では、その構成要素 (リスト) によって対話の履歴などの系内部の状態を保持する。同時に、相手からの発話や動作など、外界からのイベントはこの環境的な文脈の一部を構成する。それぞれの「ふるまい」は、実環境に対して直接影響を与え、その影響を参照すべきものであるが、インプリメント上では、この環境的な文脈が外界と系内部の「ふるまい」とをつなぐインタフェースとして機能する。

また、意図的な文脈では、その構成要素 (リスト) によって、システム全体としてのグローバルで、多重なゴールや期待などを保持する。この意図的な文脈に置かれた多重のゴールは、「ふるまい」の相互作用によって会話の流れを組織化する際に参照されるリソー

Behavior $B_i$	
precondition list:	$P_1, P_2$
add list:	$A_1$
delete list:	$D_1, D_2$
action:	action
energy:	0.0
threshold:	50.0

図2 「ふるまい」に対する記述  
Fig.2 Description for behavior.

スの1つにすぎない。スクリプトや用意されたプランの系列のように、会話の背後にあって、その流れを統率するものではない。それらがどのように行為や会話の組織化に反映されたのかは、その組織化の結果として事後的に明らかになるという性質を持っている。この環境的な文脈や意図的な文脈と「ふるまい」との相互作用については、3.4節で述べる。

### 3.3 「ふるまい (behavior)」

個々の「ふるまい」の記述は、基本的に ANA<sup>9)</sup>における記法を援用している。図2に「ふるまい (behavior)」の記述例を示す。action には、「ふるまい」が活性化された際の外界に対して表出する行為のタイプが記述されている。たとえば、ユーザに対する具体的な応答や問合せといった発話の内容が与えられる。precondition list, add list, delete list にはそれぞれ、前提条件、追加要素、削除要素がリストとして記述される。前提条件には、「ふるまい」が活性化されるために環境的文脈に充足されるべき制約が記述されている。この条件が充足された「ふるまい」を活性可能な「ふるまい」と呼ぶ。追加要素には「ふるまい」が活性化された際に、環境的な文脈に対して新たな状態として追加される要素が記述され、削除要素には「ふるまい」が活性化された際に環境的な文脈から削除される要素が記述されている。energy には、「ふるまい」が2つの文脈や他の「ふるまい」と授受する活性化エネルギーの初期値が記述され、threshold には、「ふるまい」が活性化されるために必要な活性化エネルギーの閾値が記述されている。各々の「ふるまい」は、前提条件が満足し、活性化可能な状態になっても、協調・競合の結果として集積した活性化エネルギーのレベルが閾値を上回らない限り、活性化されず外界に対して行為を表出することはできない。そのため、活性化エネルギーを媒介として、自身が活性化するために互いに協調・競合を繰り返す。すなわち、自身の持つ前提条件の充足に寄与する他の「ふるまい」に対しては、活性化エネルギーを与えることでその「ふるまい」の活性化を支援する。一方、自分の持つ前提

条件の充足を妨げる他の「ふるまい」に対しては、活性化エネルギーを奪うことでその「ふるまい」の活性化を抑制する。

### 3.4 動作アルゴリズム

インタラクティブシステムにおける、各「ふるまい」と文脈との関係、その動的な行動選択について基本的な動作を以下で説明する。

**Step 1: 相手からのイベント入力** 相手の発話や相手の身振り動作などは、イベントとして環境的な文脈に付加される。各々のイベントには、ラベルが付けられ、環境的な文脈の構成要素の1つとなる。

**Step 2: 2つの文脈からのサポート** 環境的な文脈に沿う「ふるまい」は、その文脈からサポートされる。すなわち、環境的な文脈を構成する要素のうち少なくとも1つが「ふるまい」 $B_i$ に記述された前提条件に一致する場合、環境的な文脈は「ふるまい」 $B_i$ に対して次の活性化エネルギー  $e_e(B_i)$  を与える。

$$e_e(B_i) = \frac{1}{N_e} \cdot \phi \quad (1)$$

また、意図的文脈を構成する要素のうち少なくとも1つが「ふるまい」 $B_j$ に記述された追加要素に一致する場合、意図的文脈は「ふるまい」 $B_j$ に対して次の活性化エネルギー  $e_i(B_j)$  を与える。

$$e_i(B_j) = \frac{1}{N_i} \cdot \gamma \quad (2)$$

ここで、 $N_e$  は環境的な文脈から活性化エネルギーが送られる「ふるまい」の総数であり、同様に  $N_i$  は意図的文脈から活性化エネルギーが送られる「ふるまい」の総数である。 $\phi, \gamma$  はパラメータであり、2つの文脈から「ふるまい」への影響の強さを表す。

**Step 3: 「ふるまい」どうしの協調・競合関係** 各々の「ふるまい」の間では、「ふるまい」の内部記述である3つのリスト（前提条件リスト、追加要素リスト、削除要素リスト）と環境的な文脈との関係により、以下に示す3種類のリンクが生成される。

#### 【後向きリンク】

活性化可能な「ふるまい」 $B_{i1}$  は、自身の持つ追加要素と同じ素性を前提条件に持つ「ふるまい」 $B_{j1}$  に対して後向きリンクを張る。

#### 【前向きリンク】

活性化可能でない「ふるまい」 $B_{i2}$  は、自分自身の前提条件のうち、現在の環境的な文脈に一致し

ない要素を追加要素に持つ「ふるまい」 $B_{j2}$ に対して前向きリンクを張る。

#### 【競合リンク】

「ふるまい」 $B_{i3}$ は、自分自身持つ前提条件の要素と同じ素性を削除要素に持つ「ふるまい」 $B_{j3}$ に対して1方向の競合リンクを張る。ただし、2つの「ふるまい」の間で、お互いの前提条件の要素を自分の削除要素に持つ場合には、双方向での競合リンクが張られることになる。

これらのリンクは、現在の環境的な文脈の変化に対して動的に生成される。たとえば、活性化可能でない「ふるまい」 $B_{i2}$ から「ふるまい」 $B_{j2}$ に対して前向きリンクの張られている場合に、もし環境的な文脈が変化し、「ふるまい」 $B_{i2}$ が活性化可能となることで、そのリンク関係は消滅する。同時に、「ふるまい」 $B_{i2}$ は、新たな「ふるまい」と後向きリンクの関係を持つことになる。

相互の「ふるまい」の間の関係が決定すると、この関係に基づいて活性化エネルギーの授受による「ふるまい」間の協調・競合が起こる。ここで、「ふるまい」 $B_i$ の持つ活性化エネルギーの値を $E_i$ とする。このとき「ふるまい」 $B_i$ は、3種類のリンク関係に基づいて、他の「ふるまい」に対して以下の活性化エネルギーを与える（競合リンクでは奪う）。

【後向きリンク先の「ふるまい」 $B_{j1}$ に与える活性化エネルギーの量】

$$e_b(B_i \rightarrow B_{j1}) = \frac{1}{N_{ib}} \cdot \alpha \cdot (\phi/\gamma) \cdot E_i \quad (3)$$

【前向きリンク先の「ふるまい」 $B_{j2}$ に与える活性化エネルギーの量】

$$e_f(B_i \rightarrow B_{j2}) = \frac{1}{N_{if}} \cdot \alpha \cdot E_i \quad (4)$$

【競合リンクの関係にある「ふるまい」 $B_{j3}$ から奪う活性化エネルギーの量】

$$e_c(B_i \leftarrow B_{j3}) = \frac{1}{N_{ic}} \cdot \alpha \cdot (\delta/\gamma) \cdot E_i \quad (5)$$

ここで、 $N_{ib}$ 、 $N_{if}$ 、 $N_{ic}$ は、それぞれ「ふるまい」 $B_i$ が持つ後向きリンク、前向きリンク、競合リンクの総数である。また、 $\alpha$ 、 $\delta$ はすべての「ふるまい」に共通のパラメータであり、 $\alpha$ は「ふるまい」の間での相互依存の度合いを表している。2つの文脈と他の「ふるまい」からのサポートによって、各々の「ふるまい」の持つ活性化エネルギーの値は逐次変化していく。このとき、すべての「ふるまい」の活性化エネルギー量の平均値が

一定に保たれるように、エネルギーの正規化が行われる。

**Step 4: 「ふるまい」の活性化** 活性化可能な（活性化のための条件がすべて環境的な文脈に充足されている）「ふるまい」 $B_k$ は、自身の持つ活性化エネルギー $E_k$ が閾値を超えることで活性化の候補となる。これらの候補の中で活性化エネルギーの値が最大の「ふるまい」が最終的に活性化される。ここで、閾値を上回る活性化エネルギーを持った活性化可能な「ふるまい」が存在しない場合には、活性化は起こらない。このとき、すべての「ふるまい」に対して閾値に $\beta$  ( $0 < \beta < 1$ )を乗じる。Step 4ごとに、活性化する「ふるまい」が現れるまで $\beta$ の乗算は行われ、閾値は徐々に低下していく。活性化する「ふるまい」が現れると閾値はもとの初期値にもどる。活性化された「ふるまい」は、**action**に記述された処理を実行することで外界に影響を及ぼす。同時に記述されている追加要素リスト、削除要素リストに基づいて環境的な文脈に影響を与える。

これら Step 1 から Step 4 の処理を繰り返すことによって、逐次何らかの「ふるまい」が活性化し、それらの系列によって合目的な行為が次第に組織されていく。

## 4. 音声対話システムへの実装

### 4.1 プロトタイプシステムの構成

ユーザとのかかわりにおける状況的な行為による会話の組織化を実現するために、3.1節で述べた動的な行動選択のモデルに基づいて、インタラクティブ・システムのプロトタイプを音声対話システムとして実現した。本システムは、「横浜まで、二枚、指定で」といった駅の窓口などで交わされる電車の切符販売での簡単な会話を想定している。システムは、相手の発話を逐次的に解釈しながら、相手に対する応答を合成音声により生成する。また、その過程をモニタするためにシステムが逐次的に解釈した結果をパラフレーズして画面に表示させている。図3に本システムの構成図を示す。

音声認識には、単語グラフを用いた連続音声認識システム<sup>10)</sup>を用いている。認識のための基本語彙数は155単語であり、発話の単位として「え〜」「横浜までね」「二枚」「指定で」といった発話断片 (utterance fragment) の連鎖を想定している。音声認識に対する言語的な制約は、発話断片の単語連鎖の規則としてFSA (Finite State Automaton) で記述されてお

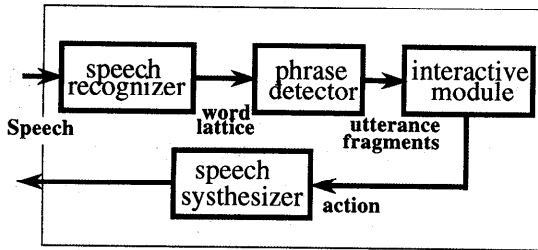


図3 システム構成図  
Fig. 3 System diagram.

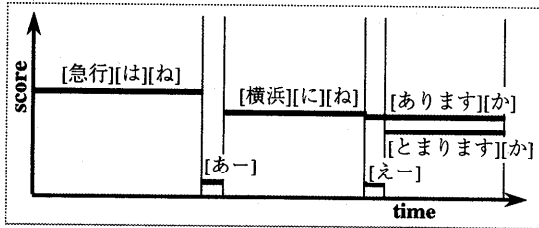


図4 発話の断片の例  
Fig. 4 Example of utterance fragments.

り、この規則で受理可能な発話断片とそれらの任意の接続を許している。phrase detector は、音声認識の結果得られた単語ラティスを FSA で記述された発話断片の単位にまとめ、その系列を出力する。図 4 に phrase detector の出力例を示す。この例では、「急行列車はね」、「あ」、「横浜にね」などが検出された発話断片であり、それらの系列はいくつかのオルタナティブを含むラティスとして出力される。これらの発話断片にはそれぞれ、素性に基づくラベルが付与される。たとえば、「横浜にね」の発話の断片に対しては、*PP.optn.destination* のラベルが付与される。この場合、文法的な素性は [任意格の後置詞句] (*PP.optn*) であり、意味的な素性は [行き先] (*destination*) に属していることを示している。これらの素性記述では、構造化を行うために各ラベルをドットで区切り、その機能を分類している。これらは各々の文脈や「ふるまい」間の協調・競合リンクを張る際に参照され、「ふるまい」の記述にあたっては正規表現なども許される。

インタラクティブモジュール (*interactive module*) は、動的な行動選択モデルに基づき、文脈や自律的な「ふるまい」との間で生まれる局所的な制約に基づいて、これらの発話断片を漸次的に解釈し、それに合わせた応答を逐次的に生成していく。

#### 4.2 「ふるまい」とその記述

プロトタイプ・システムにおいて使用されている「ふるまい (behavior)」とその記述例について述べる。「ふるまい」とは、状況に埋め込まれたエージェ

ント (*situated agent*) として実現された行為のプリミティブであり、局所的な目標に基づいて環境の状態を参照しながら、文脈を解釈し全体として合目的な行為を組織化するための基本行為、および制約の単位である。このインタラクティブモジュールで機能する「ふるまい」の総数は 127 であり、その基本的な役割により以下の 3 種類に分類される。

#### 【発話断片の間の関係づけ】

このプロトタイプシステムへの入力として想定されている発話は、FSA で定義された発話断片 (*utterance fragment*) とその連鎖である。発話断片のラティスに対して、これらの間を関係づける制約として機能する「ふるまい」が記述されている。この「ふるまい」は、制約に基づく句構造文法と同様に、句構造におけるノード間の局所的な制約により記述される。発話断片の間の関係づけを行う「ふるまい」の例を図 5 (a) に示す。たとえば、次の書き換え規則、

*SpeechAct.Actionrequest.give* →

*Pred.Actionrequest.give*, *Object.give*

に相当する「ふるまい」では、追加要素 (*add list*) に書き換え規則の左辺の素性 *SpeechAct.Actionrequest.give* (“～を下さい”) が記述される。また、前提条件 (*precondition list*) と削除要素 (*delete list*) には右辺素性 *Actionrequest.give* (“下さい”) と *Object.give* (“下さい”の目的補語) がそれぞれ記述される。このように記述された「ふるまい」において、この書き換え規則の適用が可能となるのは、「ふるまい」に記述された前提条件が環境的な文脈に満たされている場合に限られる。また、この適用の結果として、削除要素に記述された右辺の素性が削除され、新たに左辺の素性を生成する。このような発話断片間での関係づけとして記述された「ふるまい」の総数は 77 である。

#### 【時間的な重複に対する制約】

図 4 に示した例のように、音声認識の曖昧さから、音声認識結果は複数の発話断片の候補をともなったラティスの形で与えられる。ここでの時間的に重複したオルタナティブを解消するため、同一の時間軸では 2 つ以上の発話断片は存在できない制約を「ふるまい」として記述している。図 4 のラティスの例では、*Pred.YN-question.exist* (ありますか) と *Pred.YN-question.stop* (とまりますか) の 2 つの候補が時間的に重複している。この場合に発話断片を検出する phrase detector においては、それぞれの発話断片のラベルに対し、競合する認識候補が存在することを示すラベル *comp* (*competitive*) が追加される。検出された発話断片に対し、それらをまとめあげる「ふるま

**SpeechAct.Actionrequest.give**  
**precondition list:** *Pred.Actionrequest.give, Object.give*  
**add list:** *SpeechAct.Actionrequest.give*  
**delete list:** *Pred.Actionrequest.give, Object.give*  
**action:** *Actionrequest.give, object.give*  
           → *Actionrequest.give, object.give*

(a) 文法規則による制約

**dissolve time competition**  
**precondition list:** *Pred.YN-question.stop.comp,*  
                       *Pred.YN-question.exist.comp*  
**add list:** *Pred.YN-question.stop*  
**delete list:** *Pred.YN-question.stop.comp,*  
                       *Pred.YN-question.exist.comp*  
**action:** *dismiss competition*

(b) 時間的競合の解消

**trainASticket**  
**precondition list:** *\*.train*  
**add list:** *\*.ticket*  
**delete list:** *\*.train*  
**action:** *change label train → ticket*

(c) 意味素性の変更

**default.amount.ticket**  
**precondition list:** *request.amount*  
**add list:** *default.amount*  
**delete list:** *request.amount*  
**action:** *say 「一枚ですね」*

(d) 省略要素の推定(省略補間)

**fix.amount**  
**precondition list:** *default.amount*  
**add list:** *amount*  
**delete list:** *default.amount*  
**action:** *fix default amount*

(e) 推定値の確定(省略補間)

**ask.amount**  
**precondition list:** *request.amount*  
**add list:**  
**delete list:**  
**action:** *say 「何枚ですか」*

(f) 聞き返し

図5 「ふるまい」の記述例

Fig. 5 Designed behaviors and their descriptions.

い」を機能させるためには、時間的な重複を解消し、追加したラベル *.comp* を削除することが条件となる。図5(b)に、*Pred.YN-question.stop.comp* に対して時間的重複を解消するための機能する「ふるまい」の記述を示す。このオルタナティブ解消のための「ふるまい」は、*Pred.YN-question.exist.comp* に対してオルタナティブを解消する「ふるまい」との間で競合関係を持つ。これら競合する2つの「ふるまい」のうち、文脈に適合したものほど他の「ふるまい」や環境から

のサポートが強く、多くの活性化エネルギーを獲得することができる。その結果、活性化エネルギーをより多く獲得した一方の「ふるまい」が活性化されると、時間的に重複する発話断片は他の発話断片と関係づけられる機会をいったんは失うことになる。

【システムの備える意図的戦略】

「ふるまい」は、下位の要素を関係づけて、上位の要素に統合するためにボトムアップに機能するものだけとは限らない。むしろ、各々の「ふるまい」の局所的な目的を達成するために、意図的に環境的環境に働きかけ、それを書き換えていく戦略を持った「ふるまい」が、システムがより主体的に機能するために不可欠である。ここでは、目的を達成するための条件が不足している場合に、意図的に環境的文脈を作り変える戦略を備える「ふるまい」や相手に働きかけることにより、新たな発話を促す戦略を備える「ふるまい」が機能している。「ふるまい」自身が環境的な文脈の変更を促す戦略は、相手の意図を予測する行為を含んでおり、これらの相互作用によって、ユーザとシステムとの間のインタラクションはより柔軟なものとなる。

戦略的に機能する「ふるまい」の例を図5(c), (d), (e), (f)に示す。図5(c)は、「禁煙車を一枚下さい」のように、「禁煙車」が本来の意味(列車)ではなく別の意味(切符)で用いられているような場合に機能する。本来の「列車」の属性では、「下さい」の目的補語となることができない。このような場合は、その属性を戦略的に変更することで、発話を解釈する際の方略を変える。図5(d)の **default.amount.ticket** は、ある発話が省略されている場合、推定値「一枚」を環境的な文脈に戦略的に生成し、その省略を補うことで条件を満足させる機能を持つ。同時に「ふるまい」の外界に対する **action** として、「一枚ですね」との確認を相手に対して行う。この問いかけに対して、たとえば、「二枚です」との訂正が行われる場合と、そのまま聞き流される場合がある。前者の場合には、相手からの応答により、解釈の途中で環境的な文脈が変化してしまう。図5(e)に示す **fix.amount** は、**default.amount.ticket** が補った切符の枚数の推定値を確定する働きをする。たとえば、**fix.amount** が活性化する前に、相手が「二枚です」と訂正すれば、推定枚数「一枚」は環境的な文脈から削除されるため、この「ふるまい」が活性化されることはない。相手からの訂正がない場合には、**fix.amount** が活性化され、推定値の「一枚」が切符の枚数として確定する。このように、解釈の途中で、相手の発話によって環境的な文脈の変化するような場合でも、その変化に応じた解

積の方略をとることができる。図5(f)は、相手に対して必要な情報を問い合わせるために機能する「ふるまい」である。この「ふるまい」が活性化されることにより「何枚ですか」との質問を生成し、不足する枚数の情報を補うように求める。

## 5. 基本的な動作とその考察

### 5.1 システムとのインタラクションの例

図6にシステム(S)とユーザ(U)とのやりとりの簡単な例を示す。動作例1では、「下さい。大阪行きの切符を」というユーザの発話に対し、省略された切符の枚数を「一枚」と補って解釈を進めようとする過程で、ユーザから「二枚」との割り込みに対して、これまでの解釈を修正した例である。動作例2は、「喫煙車をお願いします」の“喫煙車”は“列車”そのものを意味しているのではなく、“喫煙車の切符”のことであると判断して解釈を進めている。また相手の発話を解釈する過程で生じる情報の不足に対し、相手に対する問合せによって解消している。発話の漸次的な解釈の過程で、相手からのイベントによって環境的な文脈が変化するような場合にも、新たな環境的な文脈に応じた適切な解釈を行うことができる。

### 5.2 基本動作のトレース

前述のインタラクションがどのような「ふるまい」によって、組織化されたのかを検証しておく。「喫煙車をお願いします」という相手の発話に対する活性化された「ふるまい」の系列とその制約のタイプを図7に示す。この例では、はじめに時間的な重複を解消する「ふるまい」の活性化により、音声認識によるオルタナティブが解消される。この結果を受けて、句構造規則を制約に持つ「ふるまい」が活性化され、これに関連した発話断片が漸次的にまとめあげられる。また、戦略的に機能する「ふるまい」は、省略などによる情報の欠落によって、句構造規則がそのまま適用できなくなる箇所において、適時活性化される。この戦略的な「ふるまい」の働きによって、不足する情報が環境的な文脈に補われる。こうした個々の制約、すなわち「ふるまい」が適用される手順は明示的に与えられてはいない。しかしながら、自律的に機能する複数の「ふるまい」と環境とが相互に制約しあいながら、一連の「ふるまい」がそれぞれ適切な状況で出現する。つまり、アプリアリに手続き的な適用規則を備えることなしに、入力される発話断片の系列に対して、その解釈の方略が事後的に組織化されるといえる。

この図7の例において、戦略的に機能する「ふるまい」**Oblig.trainASticket** (“列車”を“列車の切符”

インタラクション  
U:「下さい。大阪行きの、切符を」  
S:「一枚ですね。」  
U:「えーと、二枚。」  
システムの解釈  
"大阪行きの切符を二枚下さい。"

#### 動作例1

インタラクション  
U:「あの一、喫煙車を、お願いします。」  
S:「喫煙車の切符ですね。」  
S:「一枚ですね。」  
U:「はい。」  
システムの解釈  
"喫煙車の切符を一枚お願いします。"

#### 動作例2

図6 システムとのインタラクションの例  
Fig. 6 Examples of human-system interaction.

Activated behavior	Type of behavior
Compe.PP.Oblig.train	(時間的な重複解消)
Oblig.train	(文法規則)
Permrequest.give	(文法規則)
Request.Object.give	(文法規則)
Oblig.trainASticket	(システムの意図・戦略)
**喫煙車の切符ですね**	
Complement.Default.Amount.ticket	(システムの意図・戦略)
**一枚ですね**	
Fix.Amount	(システムの意図・戦略)
Add.amount.TO.Oblig.ticket	(文法規則)
Object.give.FOR.Permrequest.give	(文法規則)
SpeechAct.Permrequest.give	(文法規則)
**喫煙車の切符を一枚お願いします**	

図7 活性化された「ふるまい」の系列  
Fig. 7 Sequences of activated behaviors.

と見なす)が活性化される際の、2つの文脈と他の「ふるまい」との間の協調関係を図8に示す。図中の矢印は、活性化エネルギーの流れを示している。各「ふるまい」の記述は、前提条件をc:で、追加要素をa:で表し、他の構成要素は省略した。また、前提条件の要素でイタリックの部分は環境的文脈に存在しない要素である。この意図的文脈は、「ふるまい」**SpeechAct.Permrequest.give** (～をお願いします)に対して活性化エネルギーを与える。このとき、「お願いします」の目的補語 *Object.give* はまだ環境的文脈に存在しない。そのため、この目的補語を生成する「ふるまい」**Object.give** に活性化エネルギーの流れが集中する。ここで、この **Object.give** の前提条件には、*Oblig.ticket.amount* の記述がある。これは「お願いします」の目的補語は“xx枚の切符”であることを意味する。しかし現在の環境的文脈には、“切符”に関する要素が存在せず、**Object.give** を活性化



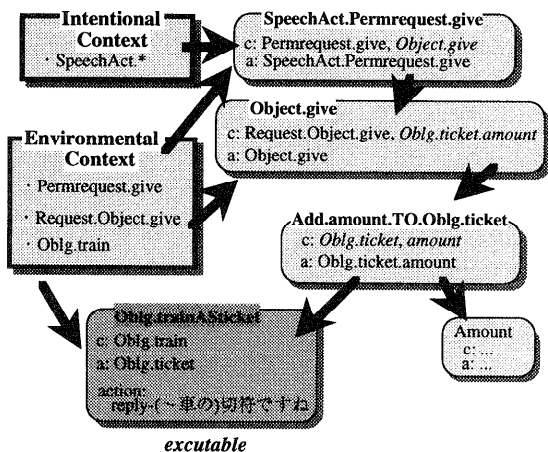


図8 「ふるまい」と文脈との協調関係

Fig. 8 Cooperative relations among behaviors and contexts.

する条件が不足している。このとき、**Object.give**の前提条件を満足させるために、*Oblig.ticket*を追加要素に持つ「ふるまい」に対して活性化エネルギーの流れが生じる。一方、環境的な文脈には *Oblig.train* (喫煙車を) が存在するため、*Oblig.train*を前提条件に持つ「ふるまい」は環境的な文脈からのサポートを受ける。これら結果として、活性化エネルギーは「ふるまい」**Oblig.trainASticket** (“列車”を“列車の切符”と見なす)に集中する。そして、この「ふるまい」が活性化されることによって、環境的な文脈に *Oblig.train*が生成される。その後、切符の枚数を補う「ふるまい」によって、*Object.give*を生成するための条件が環境的な文脈に整えられていく。

### 5.3 インタラクティブシステムにおける即応性と熟考性

ユーザとのインタラクションにおいて、省略などをもとむ相手の発話に対しては文脈に厚い、熟考的な解釈が求められる。また、急な話題の転換に対しては、即応的な解釈が求められる。

これまで述べてきた個々の「ふるまい」は、自己の局所的な目的を満足させるための行動規範を持つ。しかし同時に、「ふるまい」どうしでの相互作用は拘束条件として働くために、システム全体としては文脈を反映した行為の組織化が実現される。これらの「ふるまい」は、2つの文脈と他の「ふるまい」から影響を受ける。このバランスを変えることで、全体として性質の異なる行為を組織化させる。はじめに、環境的な文脈からの影響に比較して、「ふるまい」間での関与が小さい場合を考える。この場合には、個々の「ふるまい」の目的が優先され、環境的文脈に合致した目的

S : 「特急あずさですね」  
S : 「特急あずさの切符ですね」  
S : 「一枚ですね」

(a) 「ふるまい」間の協調関係が弱い場合 ( $\phi=40, \alpha=0.06$ )

S : 「特急乗車券ですね」  
S : 「一枚ですね」

(b) 「ふるまい」間の協調関係が強い場合 ( $\phi=40, \alpha=0.12$ )

図9 「ふるまい」の協調関係とシステムの応答 (「松本行きの特急列車をお願いします」に対する応答)

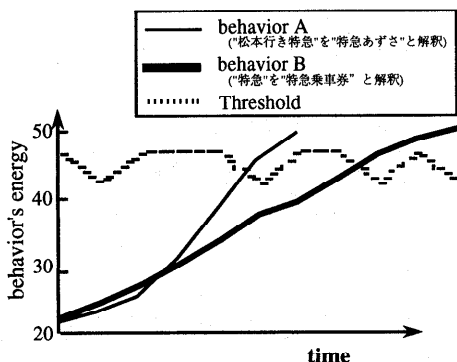
Fig. 9 Relation between parameter value and response of the system.

を持つ「ふるまい」が即応的に活性化される。このように環境的な文脈に対して即応的な反応する「ふるまい」はシステム全体の性質を即応的なものとする<sup>11)</sup>。一方、「ふるまい」間で強い相互作用がある場合には、局所的な目的よりもシステム全体の目的が優先される。つまり、システム全体の目的に合致している「ふるまい」ほど、他からの活性化エネルギーが集中し、優先的に活性化される。その結果、全体として合目的性の高い方略が選択され、それらの振舞いは熟考的なものとなる。

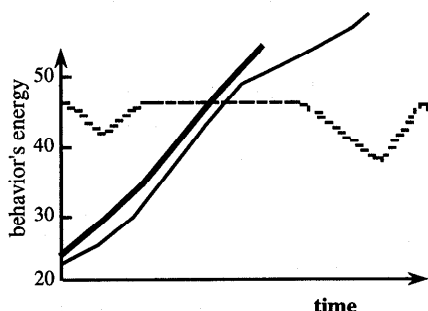
### 5.4 パラメータとシステムの挙動との関係

本プロトタイプシステムにおいて、熟考性に関するパラメータとシステムの挙動との関係を調べた。式(1)における  $\phi$  は、各「ふるまい」に対する環境的文脈からの影響の大きさを表す。また、各「ふるまい」の間での協調・競合関係の大きさは、式(3)、(4)、(5)の  $\alpha$  によって変化させることができる。これらのパラメータ  $\phi$  と  $\alpha$  に対する、「松本行きの特急列車をお願いします」という発話に対するシステムの振舞いの変化を調べた。

図9にパラメータの値の変化に対するシステムの応答例を示す。相手の発話断片を解析し、環境的な文脈に「松本行き」と「特急列車」が生成されることで、「松本行き + 列車」を「特急あずさ」と解釈する2つの「ふるまい」が活性化可能な状態になる。また、「お願いします」が環境的な文脈に加えられたことにより、「～をお願いします」という解釈を促す「ふるまい」間での協調が行われる。「ふるまい」間での相互作用が小さい ( $\alpha = 0.06$ ) 場合には、環境的な文脈に存在する「松本行き」と「特急列車」からの影響を直接受けて、「松本行き + 特急」を「特急あずさ」と解釈する「ふるまい」が最初に活性化され、遅れて「列車」を「列車の



(a) 「ふるまい」間の協調関係が弱い場合  
( $\phi=40, \alpha=0.06$ )



(b) 「ふるまい」間の協調関係が強い場合  
( $\phi=40, \alpha=0.12$ )

図10 「ふるまい」の活性化エネルギーの推移  
Fig.10 Change of behaviors' activated energy.

切符」と判断する「ふるまい」が活性化される。この際のシステムの応答が図9(a)である。一方、 $\alpha$ の値を0.12と大きくし、「ふるまい」の間での影響を強めていくと、「～をお願いします」を活性化させることを優先するために、「お願いします」の目的補語「切符を」を活性化させる「ふるまい」に対して活性化エネルギーが集中する。この結果「特急列車」を「列車の切符」と解釈する「ふるまい」が先に活性化され、次に「～行きの切符」と解釈する「ふるまい」が活性化される。この場合のシステムの応答が図9(b)である。この2つの解釈の過程を比較すると、前者は、局所的な視点から発話を漸次的に解釈しており、後者は、発話全体を文意を優先して熟考的に解釈を進めていると考えられる。

図10は、これら2つの条件において、「松本行き+特急」を「特急あずさ」と解釈する「ふるまい」Aと「列車」を「列車の切符」と解釈する「ふるまい」Bが持つ活性化エネルギーの推移を示したものである。図中、点線が活性化閾値の大きさであり、各々の「ふるまい」が活性化されるためには、この閾値を上回ることが条件となる。条件を変えた2つの結果(a)、(b)

表1 パラメータとシステムの応答との関係  
Table 1 Relation between parameters and responses from the system.

$\phi$	10				45			
	20	30	40	50	20	30	40	50
$\alpha = .06$	A	B	B	B	A	B	B	B
$\alpha = .08$	A	A	B	B	A	A	B	B
$\alpha = .10$	A	A	B	B	A	A	A	B
$\alpha = .12$	A	A	A	B	A	A	A	A
$\alpha = .14$	A	A	A	A	A	A	A	A

A = 熟考的な応答 (図9(a)の応答)  
B = 即応的な応答 (図9(b)の応答)

を比較すると、「松本行き+特急」を「特急あずさ」と解釈する「ふるまい」Aは、どちらの条件ともほぼ同じ増加傾向を示す。一方、「列車」を「列車の切符」と解釈する「ふるまい」Bは、「ふるまい」間での相互作用が強い(b)の条件の方が、この活性化エネルギーの増加傾向が急峻となる。この結果、(a)の場合では、「ふるまい」Aのレベルが先に閾値を上回り、活性化されるのに対し、(b)の場合は、他の「ふるまい」のサポートによって、「ふるまい」Bの持つ活性化エネルギーが急激に増加するため先に活性化される。

表1は、システムのパラメータ $\phi$ 、 $\alpha$ の値に対するシステムの応答傾向を示したものである。 $TH$ は活性化閾値であり、この値が小さいほど、各「ふるまい」は活性化しやすくなり、システムは即応的な傾向を持つようになる。表中、Aはシステムが図9(a)の熟考的な応答を返した場合であり、Bはシステムが図9(b)の即応的な応答を返した場合を示している。

これらの結果から、 $\phi$ を大きく、あるいは $\alpha$ を小さくすると相対的に「ふるまい」の間での影響が弱くなり即応的なBの応答が現れやすくなる。また $\phi$ を小さく、あるいは $\alpha$ を大きくすると、熟考的なAの応答が現れやすくなる。閾値 $TH$ に関しては、小さく設定することで即応的な応答が現れやすくなる傾向にある。

このように「ふるまい」間、および文脈からの影響の大きさを反映するパラメータの値を変えることにより、システムの方略を熟考性、あるいは即応性に選択することが可能となる。すなわちインタラクションの状況や資源的な制約に応じて、システムの挙動を制御することが可能といえる。

## 6. むすび

日常的な会話や相手とインタラクティブ・システムとの相互行為は、文脈における状況的な行為の相互作用によって、オリジナルに組織化される。こうした視

点からインタラクティブ・システムの設計を行う際には、各々の行為を構成する要素、すなわち「ふるまい」において、その知覚と行為とがカップリングしている点に留意すべきである。これまでのインタラクティブ・システムの設計では、相手の振舞いや発話を解釈するという行為と相手に対する動作や発話を生成するという行為は分けて考えられやすい。状況的な行為では知覚と行為とは独立ではない。この2つの行為は1つの単位を構成している。

本論文では、「ふるまい (behavior)」と呼ぶ状況に埋め込まれたエージェント (situated agent) と文脈との相互作用によって、状況的な行為を出現させることを試みた。また、この状況的な行為と文脈との相互作用によって、ユーザとインタラクティブ・システムによる問題解決行動を組織化させることを試みた。簡単なプロトタイプシステムを構成し、その基本的な動作とその性質を調べた。その結果、ここで構築された音声対話システムでは、(a) 省略をともなう発話の解釈を文脈との協調計算や相手との相互行為の結果として組織化する、また、(b) 文脈からの影響や状況的な行為 (ここでは「ふるまい」と呼ぶ) の間での相互依存の度合いを制御することにより、それぞれ性質の異なる解釈の過程をオリジナルに組織化する、などの柔軟な性質を備えることが分かった。

本プロトタイプで使用された「ふるまい」の記述は、システムの基本的な動作を検証するための暫定的なものである。一般的な領域に展開する過程では、GAを援用したクラシファイア・システムなどで使われているルール生成と選択の手法などを導入しながら、インタラクションの中で最適化していくことなどが考えられる。こうした「ふるまい」の自動生成、最適化などの検討は今後の課題である。

### 参 考 文 献

- 1) Clancey, W.J.: The Frame of Reference Problem in the Design of Intelligent Systems, *Architecture for Intelligence: The 22nd Carnegie Symposium on Cognition*, van Lehn and Newell, A. (Eds.), Lawrence Erlbaum (1989).
- 2) Simon, H.A.: *The Sciences of the Artificial*, MIT Press (1969).
- 3) Schegloff, E.A.: Repair after Next Turn: The Last Structurally Provided Defense of Intersubjectivity in Conversation, *American Journal of Sociology*, Vol.97, No.5 (1992).
- 4) Norman, D.A. and Draper, S.W. (Eds.): *User Centered System Design: New Perspectives on human computer interaction*, Lawrence Erlbaum (1986).
- 5) 岡田美智男: 口ごもるコンピュータ, 情報フロンティアシリーズ (9), 共立出版 (1995).
- 6) 佐々木正人: アフォーダンス—新しい認知の理論, 岩波科学ライブラリー (12), 岩波書店 (1994).
- 7) Clark, H.H. and Schaefer, E.F.: Contributing to Discourse, *Cognitive Science*, 13, pp.259-294 (1989).
- 8) 岡田美智男: Talking Eyes: 対話する「身体」を創る, システム/制御/情報, Vol.41, No.8, pp.33-38 (1997).
- 9) Maes, P.: Situated agents can have goals, *Robotics and Autonomous Systems*, 6, North-Holland, pp.49-70 (1990).
- 10) 清水 徹, 山本博史, 政瀧浩和, 松永昭一, 匂坂芳典: 大語彙音声認識のための単語仮説削減, 信学論, J-79-D-II, No.122, pp.2117-2124 (1996).
- 11) 山田誠二: エージェントのプランニング, 人工知能学会誌, Vol.10, No.5, pp.677-682 (1995).  
(平成10年10月15日受付)  
(平成11年2月8日採録)



小窪 浩明 (正会員)

1988年上智大学卒業。1990年同大学院理工学研究科修了。同年(株)日立製作所入社。1995~97年(株)ATR 音声翻訳通信研究所出向。現在,(株)日立製作所中央研究所研究員。音声認識, 音声言語情報処理の研究に従事。日本音響学会, 人工知能学会各会員。



匂坂 芳典

1973年早稲田大学理工学部物理学卒業。1975年同大学院修士課程修了。同年日本電信電話公社(現, NTT)武蔵野電気通信研究所入社。1986年より国際電気通信基礎技術研究所(ATR)に出向。現在, ATR 音声翻訳通信研究所, 第1研究室室長。工学博士。音声合成・音声認識を中心とした, 音声情報処理, 言語情報処理の研究に従事。日本音響学会, IEEE, 米国音響学会各会員。



鈴木 紀子 (正会員)

1992年横浜国立大学工学部電子情報工学科卒業。1994年同大学院工学研究科電子情報工学専攻修士課程修了。同年、ATR通信システム研究所入所。現在、ATR知能映像通信研究所研究員。ヒューマン・コンピュータインタラクションの研究に従事。電子情報通信学会、日本音響学会、人工知能学会各会員。



岡田美智男 (正会員)

1982年宇都宮大学工学部電子工学科卒業。1987年東北大学大学院工学研究科博士課程修了。同年、NTT基礎研究所情報科学研究部。1995年よりATR知能映像通信研究所、主任研究員。工学博士。日常でのコミュニケーション場面における「身体」の主體的な振舞いとその環境、他者、メディアとのかかわりに興味を持つ。著書に「口ごもるコンピュータ」(共立出版)等。認知科学会、人工知能学会、ソフトウェア科学会等会員。