

「対話のありうる姿」 - 対話現象に対する構成論的アプローチ -

猪口聖司 鈴木紀子 岡田美智男

ATR 知能映像通信研究所

〒619-02 京都府相楽郡精華町光台2丁目2番地

{inokuchi, noriko, okada}@mic.atr.co.jp

あらまし:

二人の、あるいは多数の自律したエージェント同士が勝手におしゃべりをはじめたらいったいどんな会話をするのだろう、どんな話題が展開されていくのだろう。あるいは彼(彼女)らの相互作用の中からどんな社会が創り出されていくのだろう。興味は尽きない。そして自律エージェントにとっての「自律性」や「個性」とは何か、「社会性」とは何か。こうした「自律したエージェントの相互作用の中から、日常のおしゃべりを創り出す」という試みは、個人的な興味以上のさまざまな議論のきっかけを与えてくれるのではないだろうか。あるいは、われわれ自身の「対話のありうる姿」を教えてはくれないだろうか。本稿では、二人の自律したエージェントの相互作用によって「日常のおしゃべりを創り出す」という対話現象に対する構成論的なアプローチの概要と、われわれの「対話のありうる姿」を明らかにする試みの一端を紹介したい。

Dialogue as it can be

-A Constructive Approach to Phenomena in Dialogue-

Seiji Inokuchi Noriko Suzuki and Michio Okada

ATR Media Integration & Communications Laboratories

2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto-fu, 619-02 JAPAN

{inokuchi, noriko, okada}@mic.atr.co.jp

Abstract:

When two or more autonomous agents start chatting on their own willfulness, what kind of conversation will emerge? What kind of topic will develop? Or, what kind of society will be constructed from their interaction? There are so much things to think of. What is 'autonomy' or 'personality' for autonomous agents? What is 'sociality' for them? Beyond the personal interest, the approach of 'constructing daily chat or conversation from the interaction of autonomous agents' may give us the chance of good discussion, we think. Or, it may tell us the style of 'dialogue as it can be'. In this paper, we outline our project to find the 'dialogue as it can be', and the constructive approach that 'create the daily chat or conversation' from the interaction of two autonomous agents.

1 はじめに

「日常のおしゃべりを創り出す」ということを考えてみると、それは実に奥が深いことに気づく。例えば、統計的な言語モデルをふたつ並べて、相互作用をさせても、新たな「おしゃべりを創り出す」ことにつながらない。外界の刺激に対して線形に反応するCGとの相互作用を試みると、すぐに「飽き」を感じ、そこから新たな対話は生まれてこない。また、これまでの多くの対話システムとの対話では、システムに用意されているスロットを埋めることに終始し、新たな話題に移行していかない。おしゃべりそのものが新たに創り出されていかない。日常でのおしゃべりや対話を組織化する根源的な原理を明らかにしていく必要がある。こうした議論は、われわれの備える「自律性」や「社会性」などに関連しており、「おしゃべりを創り出す」という試みの中から、日常の対話を行なわれわれ対話者が備えるべき本質を明らかにできるものと考えている。

本稿では、特に対話の過程を「自律的な情報の生成機構」として、また、「日常でのおしゃべり」におけるさまざまな挙動を創発現象として捉え直す。そして、さまざまな要素(マルチエージェント)の相互作用から対話に特有の挙動を創り出し、その過程の中で対話のメカニズムを明らかにしたいと考えている。これは、ある意味で対話現象に対する構成論的アプローチであるといえる。

2 日常のおしゃべりに対する構成論的アプローチ

日常のおしゃべりをどのように特徴づけていけばよいのだろうか。対話研究では、他の研究分野と同様に研究領域の極度の細分化が進んでいる。例えば、発話の生成系と発話の解釈系とは独立したコミュニティで議論されることがある。また、話し手と聞き手の関係で対話者を二分して捉えてしまいがちである。その間でやり取りされる情報は、当然のように「コードモデル」の枠組で捉えられる。これまで発話は「意図された行為(intended actions)」であるという視点で捉えられ、定式化が進められてきた。ある事柄を伝える前に、伝えるべき内容が用意されている。送られてくるコードを解釈する際に、解釈すべきコードが存在している。また、幾つかの処理モジュールを仮定し、その中で情報の流れを考える。いわゆる「情報処理アプローチ」と呼ばれる枠組である。このアプローチでは、行為の遂行に先だって、その行為をナビゲートするプランなどの情報が用意されていると考える。そして、モジュール間を流れる情報は、一般化されていることが期待される。例えば、対話システムでの「意味解析部」は、解釈の目的を知らされないまま、一般的な意味の解釈を求められる。これら「コードモデル」や「情報処理アプローチ」に共通するのは、いわゆる

「二元論」的な説明である。しかしながら、自律したエージェントを実世界で動かすことを考えると明らかであるように、行動主体(エージェント)にとっての外界の知覚は、その行為と独立には議論できない。

「わたしたちは動くために知覚するが、知覚するためにはまた動かなければならない」、こうしたギブソン(J.J.Gibson)の言を待つまでもないだろう。個人差もあるだろうが、日頃ほとんど留守番電話を利用しない人も多いという。留守番電話であると感じたとき、伝言を託すことを諦め、受話器を置いてしまう。何かを伝えようとしたわけではない。むしろ、本当に伝えたいことは、おしゃべりの中で生まれてくることなのかもしれない。「コードモデル」を仮定してデザインされた留守番電話に対してそうであるように、日常の何気ないところに様々な違和感を見出すことができる[1]。

われわれはことばや対話、コミュニケーションに対しても、何かあらかじめ与えられた構造や背後にそれを操っているものの存在を仮定しがちである。ところが日常での会話では、司会者などをとくに必要とすることなく、発話権の交替が無意識に行なわれ、次々に話題が生起し、展開していく。これらの振る舞いや制御方略そのものは、さまざまな要素の局所的な相互作用の中から、結果として立ち現れたもので、制御が系の外から与えられたものではない。こうした日常のおしゃべりに共通することは、行為そのものが状況の中で組み立てられ、レギュレートされることだろう。

ここでは、ことばや対話は本質的に「創発的な性質(emergent property)」を有するものであるとの視点から、「日常のおしゃべり」に代表される、われわれの日常での振る舞い、無意識な、あるいは非恣意的な振る舞い、さらにはノンバーバルな情報の扱いを考えよう。多くの分析的な研究がそうであるように、それぞれの要素を他の要素(すなわち、その環境)と分断した状態で眺めていては、その挙動を創り出す本質的なメカニズムやダイナミクスを明らかにするのは難しいように思われる。むしろ、不要な要素を削り落して単純なメカニズムとしながらも、会話の挙動をできるだけ「総体」として捉えていく必要がある。そのためには、実際に相互作用する全体の系を構成し、そこから湧きあがってくる挙動を観察するという手段をとる必要がある。

3 状況の中で組織化される行為

日常の発話を「創発的な挙動」として捉える立場では、発話(すなわち行為)に先だって、その発話をナビゲートするプランの存在を仮定しない。むしろ、環境との相互作用の中で、行為が引き出され目的的な発話そのものが結果として組織化されたと考えられる。これまでの「発話を意図された行為(intended actions)」と捉える立場をトップダウンの過程とすれば、むしろボトムアップな過程である。また、「意図

された行為」に対して、「環境から引き出された行為、環境に埋め込まれた行為 (situated actions)」と呼べるだろう。失語症などの神経心理学の分野においても、意図的な能力と自動的な能力とが解離していることが知られており、バイヤルジェ-ジャクソンの法則 (自動性-意図性解離の法則) と呼ばれている。このように、意図的に創り出される発話と状況に引き出されるかのように、無意識にでてくる発話とを区別して考えることができる。

この「状況に引き出される発話」のメカニズムを考える上で、二つの環境の存在を仮定している。一つは、対話の状況を反映するもの (Environmental Context) であり、もう一つは、話し手の意図や動機を反映するもの (Intentional Context) である。われわれの抱く意図や動機も行動選択における環境 (リソース) にすぎないと考える。ある環境の下で出現した行為は、新たな環境を作りだし、次の行為を導き出す。この連鎖的な循環は、流暢な振る舞いを創り出す。これを系全体として見ると、系の挙動の中から湧きあがってきた行為が新たな環境を形成し、その環境が新たな行為を湧きあがらせるための境界条件となっている。こうした循環は「創発過程」そのものである。この行為を組織化する過程により、あらかじめ抱いていた意図や動機が反映され、合目的な行為へと導かれる。また、常に環境に対する「行為の可能性」の知覚を伴う。このことは行動主体の外に対しても同様であり、外界のイベントの知覚は行為と独立ではない。手振りや視線の動きにも、恣意的なものや非恣意的なものがある。その幾つかは、社会的なコードが与えられており、われわれはそのコードを解釈することでコミュニケーションを成り立たせている。しかし、他の多くは社会的なコードの与えられていない非恣意的なものである。こうした現象は扱いにくく、一般には何らかの社会的なコードからは逸脱したもの、扱えきれない「ノイズ」と見なされることが多い。無意識な振る舞いを扱う枠組が欠如しているのである。ここでは、手振りなどのノンバーバルな振る舞いも状況から引き出され、合目的な行為を組織化する過程で、その一部を担うものと捉えている。

対話の過程に対して、こうした「状況に引き出された行為」が重なり合い、合目的な行為をボトムアップに組織化するという視点で捉え直すと (図1)、「日常でのおしゃべり」の幾つか面白い性質が明らかになってくる。その詳細については、本稿の後半で考察を進めていく。これまでのコミュニケーションがバーバルな情報を中心に「コードモデル」に基づいて偏った扱いを受けているとすれば、コミュニケーションにとってもっとも根源的な部分がすっぽり欠けているのではないかと思われる。

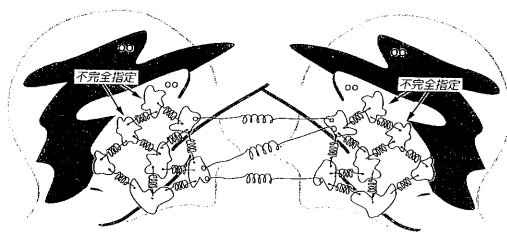


図1: 対話の組織化

4 エージェント (= 「ふるまい」) の集合と動的環境

日常なおしゃべりにおいては、発話はさまざまな「ふるまい」の中から創発する。このような日常なおしゃべりをモデル化するためには、どのような記述がふさわしいのだろうか。「ふるまい」は動的に変化する対話という環境のなかでさまざまな相互作用を行ない、ある時は環境に適応的な振舞いが、またあるときは意図に忠実な振舞いが選ばれる。いわば、対話は動的な環境の下で振舞いを選ぶことが繰り返されながら進行していくのである。このようなメカニズムをうまく記述するため、本研究では Agent Network Architecture (ANA) [2] をもとにした動的選択モデルを用いている。ANA は多数の「ふるまい」の相互作用のなかから、動的に変化する環境にふさわしい行動選択が創発するという特徴を備えており、われわれの「対話観」を記述するのにふさわしいモデルである。

ANA の基本単位 = 「ふるまい」は、環境から影響をうけて活性化される。環境は状況や目的など、対話において影響を及ぼす要素により構成されている。「ふるまい」の間には協調や競合の関係が存在し、活性化された「ふるまい」は他の「ふるまい」との関係にしたがって、お互いに活性や抑制をする。こうして環境の状態に応じて「ふるまい」が選択され、行為として遂行される。行為は環境の状態を変化させ、つぎの「ふるまい」の選択に影響を与える。このように「ふるまい」と環境との相互作用や「ふるまい」どうしの相互作用のなかから、環境の状態に応じた行為が創発する。

4.1 対話におけるマルチモダリティの表出

われわれが最初に構築したコンピュータ・エージェントは、トーキング・アイ (図2) とよばれている。トーキング・アイは二つのモダリティ、音声と動作をあやつることができる。

トーキング・アイの動作は環境と意図の間のせめぎあいから創発する。このメカニズムを ANA による動的選択で実現している。うなずきや質問、同意や否定

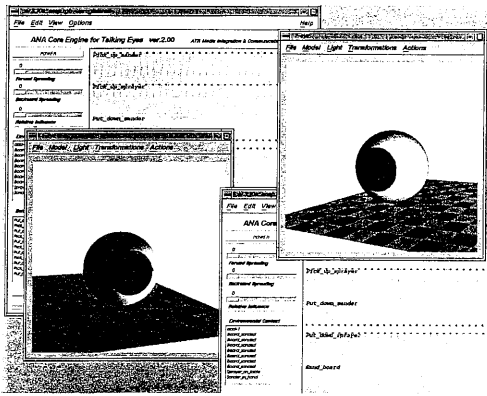


図 2: トーキング・アイの外観

など、対話における基本的な行為の単位を ANA における基本単位 = 「ふるまい」として記述し、それぞれの行為は音声や動作として創発する (図 3)。

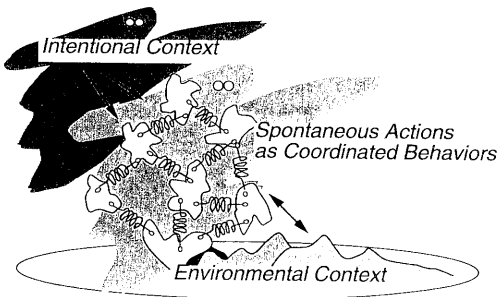


図 3: 環境と意図による行為の創発

行為が実際に行なわれると、環境が影響を受け、変化化する。変化した環境はそれまでとは違った影響をトーキング・アイに与える。トーキング・アイの動的選択のメカニズムは環境の変化をうけて、それにふさわしい行為を新たに選択する。このようにして対話における行為と環境の間で相互に影響が循環し、対話が進行するのである。また、環境には対話の相手も含まれる。対話の相手もまた環境から影響を受け、行為を選択し、環境を変化させる。このようにして、対話をおこなう両者が環境を共有しながら、対話が進んでいく。

4.2 マルチモダリティの解釈の創発

対話においては、さまざまなノンバーバルなイベントが発生している。対話者は単に対話の言語的な情報 (バーバルなイベント) のみを対象にして対話を進めているのではない。雄弁に語るひとを思いうかべるとき、直立不動で一点を凝視するような硬直したイメージを思い浮かべる人は少ないだろう。むしろ、身振り

手ぶりや生き生きとした表情、豊かなイントネーションをつけた活力にあふれる声などを自在にあやつるひとを思い浮かべるにちがいない。この「生き生きとした」という表現に対応するのは言葉使いなどのバーバルなイベントではなく、どちらかというとノンバーバルな情報である。おそらく、人間は対話においていろいろなノンバーバル情報を使って、言葉だけでは表現の難しい、豊かなコミュニケーションを行なっていると思われる。

ノンバーバル情報を含めた対話におけるマルチモダリティを解釈することも、要素の複雑な相互作用から産まれてくるととらえることができる。一般に解釈がひとつに落ち着くことはめずらしく、とくにマルチモダリティのようにいろいろな情報を統合するような微妙な解釈の場合には、いろいろな解釈がつねに共存していると考えの方が自然である。解釈が複数ある理由は、個々のモダリティの断片的な情報をもとに微妙な解釈がすこしずつ形成されてゆくために、つねに解釈の候補が存在するからであると考えられる。この場合、断片的な解釈という振舞いが、与えられた情報によって形成される環境の中でどのように統合的な解釈を形成するという過程そのものが創発的であると考えられる。

5 様々なシステムへの応用

3章では、日常のおしゃべりを創り出すメカニズムを持つエージェント間における対話のあるべき姿について述べた。しかし、このような構成論的アプローチは、トイ・ワールドでの議論に陥りやすい。それを回避するため、本章ではエージェントと実世界との関係について考えていく。5.1 では提案した日常のおしゃべりを創り出すメカニズムの実世界への応用の一つとして、現在構築中の人間とヘテロジニアスな自律エージェントとのインタラクティブシステムを紹介し、5.2 では実世界での人間のコミュニケーションのパートナーとしてエージェントが備えるべき性質について述べる。

5.1 「複数の心」との対話

人間と対話システムとのインタラクションにおいて、人間がシステムに対してやさしく親しみを感じるためには、どのようなメカニズムが必要だろうか。

人間同士の日常的な対話はさまざまな要因が関与し複雑な挙動を示す。観測された対話の現象を分析したり、解析しようとする、その挙動の多様さに圧倒される。しかし、「サイモンの蟻」[3]の議論にもあるように、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。単純なメカニズムを相互作用させることにより、複雑な挙動は容易に創り出すことができる。これらの日常的な対話や自然な発話の挙動を捉えるためには、これまでの対話システムのように、一方的に流れてくる現象を傍観者として

観察するのではなく、参加者としてむしろ現象を創り出す必要がある。ここでは「おしゃべりの場」を創り出し、それを共有するという視点から、人間と自律したエージェントとの日常でのおしゃべりを捉える。

また、これまでの対話システムでは、人間とシステムが一对一でコミュニケーションを行っていたため、常に相手の視線に監視され、自分も相手の一挙一動に注意を払わなくてはならず、人間がシステムに対してある種の切迫感を感じることがあった。これはある意味で非日常的なコミュニケーションであるといえる。そこで、人間とシステムの間人間同士に見られるような日常のおしゃべりの場を創り出すためには、人間のコミュニケーションのメカニズムをシステムに持たせる必要があると考えられる。そのメカニズムとして、人間の心は複数の小さな心で構成されるという「複数の心」の考え方 [4] を利用し、各々が日常のおしゃべりを創り出すアーキテクチャを持つヘテロジニアスなエージェントを用いて、「複数の心」が人間と対話を行うインタラクティブシステムを現在構築している (図4)[5]。

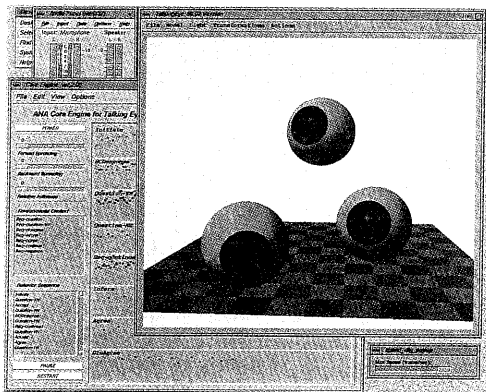


図4: 「複数の心」の自律エージェントの姿

さらに、人間と「複数の心」がインタラクションを行うことについて、次のような利点があげられる。

「社会への参加としてのインタフェース」: 「複数の心」がそれ自体一つの社会を構成していると捉えることができるため、人間がその社会に必要なときに自由に参加できるインタフェースの形態の実現が可能となる。

「インタフェースにおける他者の役割」: 人間が一つの心と対話を行っている間、他の心はその対話の成り行きを客観的にながめることができる。コミュニケーションの中で他者として存在することが可能である。

「インタラクションにおける不完全指定」: 「複数の心」同士の対話に人間が必要ときに口をはさむこと

ができる。次の発話の内容を人間が逐次指定しなくても、部分的な指定を行うことによってある意味で勝手におしゃべりが進んでいく。

これらの理由から、人間が「複数の心」とインタラクションを行うことによって始めて、誰もが参加しやすいコミュニケーションが実現できるといえる。

5.2 おしゃべりの相手としてエージェントに必要な性質

われわれは日頃、誰とおしゃべりをしているのだろうか。それは「壁」ではないだろう。おしゃべりの相手としてのエージェントの自律性、社会性の有無は、「おしゃべりの場をボトムアップに創り出せるか否か」にかかっている。例えば、動画像のように動きはあってもまったく自律性のないオブジェクトを相手にしては、おしゃべりの場は作れない。同様に、外界からの刺激に対して線形に反応をするCGを相手におしゃべりの場を創り出すことはできない。これまでの対話システムの研究では、日常のおしゃべりの場を創り出し、それを共有するという視点に欠けていた。臨場感会議システムなどでは、自律性を持たないオブジェクトが対話参加者のコミュニケーションに介在する。また、テレビ電話などでは、相手の動きを反映する動画像である。これらは、動画像やCGのオブジェクトの動作そのものを完全に指定するため、時間的な制約を受けないおしゃべりは実現できない。

われわれは「エージェントの自律性」という視点から、さまざまなコミュニケーション環境、コミュニケーションの形態を考えようとしている。エージェントの自律性を本稿での対話におけるインタラクションの文脈で議論するならば、「何らかの動機を創り出し、外界との相互作用の中で問題を解決する能力」として捉えることができる。完全に自律したエージェントの対話では、外界の人間が関与せずに勝手におしゃべりはじめてしまう。

ここでのシステムに使用した状況に埋め込まれたエージェント (situated agent) は、正確には「半自律エージェント」である。まだ自らゴールを創り出すことはない。不完全指定の形で、部分的に意図や動機を与えておくと、外界との相互作用の中で問題解決を行なう。こうした「半自律エージェント」を介在させ、コミュニケーションに不完全指定 (under-specification) の概念を導入することで、ある意味で時間的な制約から解放されたコミュニケーション (おしゃべり) を実現することができる。このような時間的な束縛を受けないコミュニケーションの例として、図5に示すような WWW (World Wide Web) への応用があげられる。対話の参加者は、自分の WWW ホームページ上で、エージェントの性格付け、多重のゴールを宣言しておき、その指定を反映させて、不特定多数のおしゃべりを実現する。当人の不在の間に

部分的に問題を解決することができる。このプロジェクト (“Talking Eyes on the Web”) の詳細は、別の機会に報告したい。

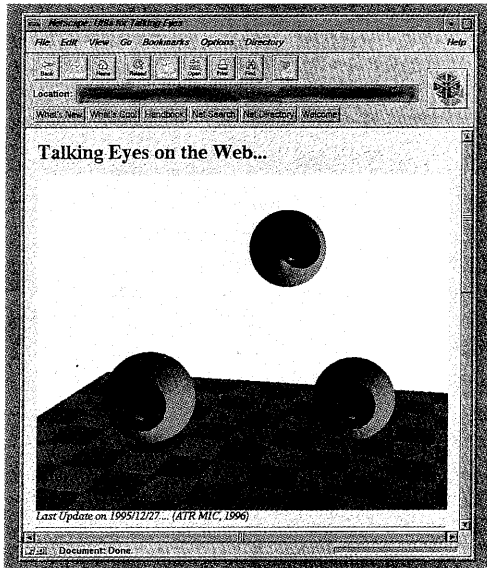


図5: 「複数の心」のWWWへの応用

「エージェントの社会性」とは何だろうか。本稿では、「実世界の人間とのインタラクションの中でひとつのシステムを形成していく能力」をエージェントの社会性として捉えることができる。日常生活において、われわれは新しい「ニッチ (生態学的位置) [6]」を獲得しながら、様々な「ニッチ」に囲まれて生活している。われわれの生み出す「人工物」は、そこに棲む人間にとっての「ニッチ」として、獲得されなければならない。「人工物」がわれわれとの間で「ひとつのシステム」を形成する能力が備わっていなければ、実世界の人間と人工物であるシステムとの間にミスマッチが生じ、ある意味で人工物が実世界から排斥されてしまうだろう。

このように生態学的な視点を用いて捉え直した自律性や社会性を持つことによって、エージェントが実世界の人間のコミュニケーションのパートナーになることが可能になると思われる。

6 おわりに

二人の自律した行為主体 (An Autonomous Agent) の相互作用の中から「日常のおしゃべりを創り出す」という観点で進めている。われわれの対話研究に対する構成論的なアプローチを紹介し、現象に対する簡単な考察を進めた。われわれの日常的な会話は、状況に引き出された発話から成り立っている。日常での会話の自然さや流暢さを支えると同時に、それは結

果として「対話的な関わりを維持する」という合目的な行為を形作っている。つまり、おしゃべりそのものの、そしてその秩序は共同行為の中でボトムアップに組織化されると考えられる。さらに、対話者間のカップリングの要素は発話に限られるものではない。ノンバーバルな情報源である手振りや「うなずき」は、状況に引き出され、新たな状況を創り出す。それらのイベントは、多様な状況に依存して解釈され、次の行為を引き出す。これらの循環的な関係は「対話者間でひとつのシステム」を構成するために重要な役割を果たしている。そして対話者間での情報伝達はどうであろうか。それはもはや事前に用意した情報をコード化して伝えるという「コードモデル」で説明される情報伝達のメカニズムではない。むしろ「相互作用の中から創発した情報を共有する」という視点で見ていく必要がある。

「ゆきあたりばったり」という言葉には、「主体性のない、いい加減な行動」という語感を伴う。しかしながら、今後もうすこし真剣に「ゆきあたりばったり」な対話をするシステムのメカニズムを議論していく必要があるだろう。最後に紹介した「複数の心」との対話に関する研究は、その一つのアプローチでもある。今後、モデルの精緻化を図り、「うなずき」、「発話権交替」などの対話現象の創発性の考察をさらに進めていく予定である。

謝辞

日頃議論して頂く (株) ATR 知能映像通信研究所 中津良平社長、第4研究室 片桐恭弘室長、立命館大学工学部 羽尻 公一郎氏に感謝致します。

参考文献

- [1] 岡田美智男: 『口ごもるコンピュータ』 情処学会編、情報フロンティアシリーズ 共立出版 (Aug. 1995)
- [2] Maes, P.: “The Agent Network Architecture (ANA)”, SIGART Bulletin, Vol. 2, No. 4, (Aug. 1991)
- [3] Simon, H.A.: “The Sciences of the Artificial”, MIT Press (1969)
- [4] Minsky 著, 安西 訳: 『心の社会』 産業図書 (Nov. 1990)
- [5] 鈴木, 猪口, 岡田: “「複数の心との対話」 - 日常的なコミュニケーションの場の創出 -” 情報処理学会第52回全国大会, 発表予定 (Mar. 1996)
- [6] J.J.Gibson 著, 古崎 他 訳: 『生態学的視覚論』 サイエンス社 (Apr. 1985)