

インタフェース・エージェントに関する基礎検討

間瀬 健二 シドニー・フェルス 江谷 為之 アルミン・ブルーダリン
{mase,fels,etani,armin}@mic.atr.co.jp

ATR 知能映像通信研究所

あらまし エージェントについてはいろいろな立場での研究があり、その定義も多様である。ここではマクロな立場でエージェントをとらえて、インタフェース・エージェントを定義して議論を進める。まずインタフェース・エージェントが持つべき機能を概観し、エージェントとのインタラクションおよびそのデザインについて考察する。現在設計を進めているエージェントシステムのベースとなる非同期階層型エージェント・アーキテクチャおよび、擬人性を表現するためのキャラクタ生成手法を紹介する。

A Basic Study on Interface Agents

Kenji Mase Sidney Fels Tameyuki Etani Armin Bruderlin
{mase,fels,etani,armin}@mic.atr.co.jp

ATR Media Integration & Communications Research Laboratories

Abstract In this paper, we define the interface agents as “autonomous computer program which has life-likeness, performs some task or tasks on behalf of an entity, communicates with other entities including humans in the world and responds to activities which occur in the agent’s internal world or the external world” and discuss on it. First, we summarize the desired functions of interface agents and consider interaction with agents and its design. We will introduce the Asynchronous-Hierarchical Agent (A-HA) Architecture which would become the base structure of our agent development and character creation method to provide a life-likeness to agents.

1 はじめに

人間同士のコミュニケーションにおいては、協同で作業をしたり作業の依頼や命令という中心的なタスクを実行する際に、創造的発想、相互理解、自己理解、話題の整理あるいは話題の展開などの重要な側面がある。我々は、コミュニケーションにコンピュータが介在しこのような側面を支援する際に使われるツールやインタフェースの研究をしている。本文では、我々が現在進めている議論活性化支援エージェントやガイドエージェントの研究プロジェクトを紹介しながら、エージェントとインタフェースの設計における課題について検討する。

エージェントについてはいろいろな立場での研究があり、その定義も多様である [1]。ここではマクロな立場でエージェントをとらえ、インタフェース・エージェントを「人間やそのほかのもの (entity) のためのあるタスクをもつ自動プログラムで、自律性があり、life-likeness (以下、生きいき性と訳す) を有して、ユーザである人間とのインタフェースまでかかわるソフトウェアなどの概念」と定義

しておく。例えば我々の考えるエージェントとは、ミュージアムのガイドエージェント [2] や議論支援エージェント [3][4] とよぶものであったり、ゲームのプレイヤーの一人になるようなエージェントである。人間同士あるいは他のエージェントとの非同期的なあるいは同期的なコミュニケーションに介在してコミュニケーションを支援するというタスクをもち、ユーザとのインタラクションの中から情報を獲得して、状況を把握し、自律的に情報フィルタリング [5] をしたりアクションを起こし、コミュニケーションを活性化させる情報を提供したりする。また擬人性を有してノンバーバルな表現で作業内容や状態を可視化したりする [6]。これは、文献 [1] で分類されている応用レベルの用法 (Interface Agent や Believable Agent) にあたると考えてよい。

以下、まずインタフェース・エージェントが持つべき機能を概観し、エージェントとのインタラクションおよびそのデザインについて考察する。現在設計を進めているエージェントシステムのベースとなる非同期階層型エージェント・アーキテクチャおよび、擬人性を表現するためのキャラクタ生成手法

を紹介する。個々のエージェントの内容については本文では詳しく触れないで概略にとどめる。

2 インタフェース・エージェントの機能

インタフェース・エージェントに要求される機能は何であろうか。石田 [1] によれば、エージェントに要求される技術は、自律性、知性、コミュニケーション、擬人性、パーソナリティなどである。我々はインタフェースにエージェント的考え方を持ち込むことによって、さらに次のような側面が強調されると考える。

- エージェントに対する認知性：振舞い (behavior) の予測や作業 (action) の可視化が行われ、エージェントの作業や振舞いに対する認知性がある。擬人性 (anthropomorphy) や生きいき性 (life-likeness) が助けになると思われる。
- エージェントの操作性、エージェント空間の身体性：大規模な情報検索、パターンマッチング、情報提示などに於ける膨大なデータを直接見せられても人間であるわれわれは操作できないし処理しきれなくなってしまう。しかしエージェントを介在することによって、これらの情報処理指示、過程および結果の操作が容易になる。たとえば、エージェントにメタファー的あるいはアイコン的なジェスチャ [7] をさせることによって、メッセージの理解や記憶が容易になる。
- エージェント行動の社会性、適応性：エージェントは知性を備えて人間とのインタラクションに対応する必要がある。例えば暗黙的な動作や状況を察知して、それに合わせた対応をするための動作認識理解を必要とする。また、個々の状況や個別の人間の性格にあわせて適応的に行動するためには学習能力が不可欠である。
- 親密性：ツールとして親密性が上がり、慣れ親しんだり、操作性を向上させ、創造性を刺激することは重要な要素である。
- 自律性：エージェントは指示された範囲での自律的な活動が認められ、そのための知性を有することが望まれる。自律的な活動により、ユーザが煩わしきから開放される可能性がある。

このように、インタフェース・エージェントを実現することによって多くの恩恵を受けることができる可能性がある。その実現には多くの課題があるが、鍵となるのは、擬人性の付加、作業空間の可

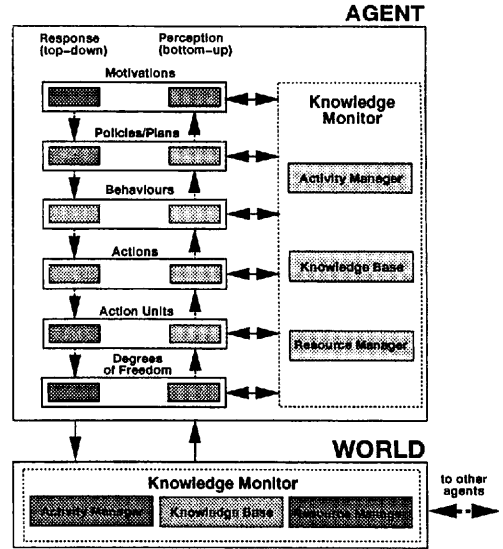


図 1: 非同期階層型エージェントアーキテクチャ

視化、状況理解、個人適応などであり、それらの制御である。また、個人適応されたエージェントなど別々のタスクをもつエージェントがお互いにコミュニケーションする際のメカニズムが必要である。

一方、知的エージェントについては数々の欠点が指摘されている。例えば直接操作性が失われたり、エージェントの作業への信頼性を確保できないなどである。われわれは、上記のような機能をもつエージェントを設計することによってある程度この問題は解決できると考えている。またそれと同時に、エージェントのアクションにいろいろなレベルで介入して直接操作ができるようなバスを提供することが必要であると考え。一方、知性と擬人化程度のバランスは長く見過ごされて、リアルな顔の表示 [8] などが研究されてきたが、最近はこの問題もよく指摘され議論されている。犬や鳥などの動物キャラクター [9][10] や幼児キャラクター [11] が受け入れられやすいのは、過度の知的インタラクションを期待させないことがメリットであると考えられている。

3 エージェントとのインタラクション

エージェントとのインタラクションをどうデザインするかは、エージェントの振舞いや行動のデザインに関わる。エージェントに自律性をもたせるためにはその行動を生成するメカニズムが重要となる。エージェントの振舞いや行動は goal-based と script-based の 2 種類が考えられる。goal-based なシステムは途中の振舞いや外部刺激に対する反応が

エージェントに任せられるため直接操作性が低い、自律的な振舞いのなかから当初予想しなかった結果をもたらしてくれる可能性がある。一方 script-based なシステムではユーザの SCRIPT が支配するためエージェントの自由度は低くなり、ユーザも SCRIPT の範囲で振る舞うことが要求される。しかしながら、エージェントの作業や行動への信頼性はあがる。どちらも一長一短があり、どちらも使えるようにするのが理想である。

以下、それぞれの方法について例をしめす。

3.1 Goal-based Behavior

Goal-based なエージェントの振舞いモデルを用いたアーキテクチャとして我々は非同期階層型エージェントアーキテクチャ (Asynchronous-Hierarchical Agent Architecture, A-HA アーキテクチャ) を構築中である。それは図1のように各エージェントはボトムアップの認識系とトップダウンの行動系からなり認識結果が行動系をドライブする。それぞれの系は上から動機 (motivation)、計画あるいはポリシー (plan or policy)、振舞い (behavior)、行動 (action)、行動単位 (action unit) というぐあいに階層的になっている。例えばトップダウンの行動系では、動機またはゴールが与えられ、ゴールへ到達するための全体の計画をきめ、計画の各段階での振舞いを決定し、振舞いにしたがって行動する、というように詳細化される。それぞれのモジュール間でのデータのやりとりはフィルタリングされ、入力にしたがってあらかじめ与えた関数やテーブル化されたリストから計画を立てたり行動をきめ、実際に行動に移す。

また A-HA はマルチエージェントアーキテクチャになっており、世界 (world) を介してコミュニケーションをとり、お互いが相手の状況を認識し、その入力をもとに自分の行動決定過程を修正したり反応する。また、各エージェントは知識ベースマネージャ、アクティビティ・マネージャ、リソース・マネージャを有し、観測世界の個別コピーを管理したり、アクション環境を管理する。

ユーザは3つの方法でエージェントの世界に介入することができる。第1は、ひとつのエージェントとなって世界でアクションを起こすのである。他のエージェントがそれを観測、認識して行動を起こす。ユーザが認識できるのも他のエージェントが世界で起こしている行動だけである。エージェントの内部状態を知るためには各エージェントがそれを世界で表象化するように設計する必要がある。各エージェントは世界の認識系にユーザである人間エー

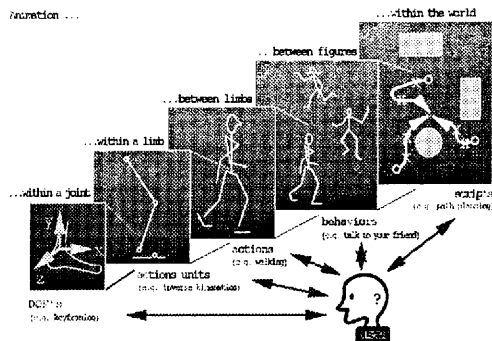


図2: 階層型の擬人化エージェントキャラクタデザイン

ジェントの行動理解センサを必要とする。ゲームでいえば、ユーザと各エージェントがそれぞれプレイヤーとなってゲームを進めるようなものである。第2は他と同じように、トップダウンに各エージェントのゴールや動機を一旦は与えるが、あとはエージェントが世界で起こす行動を観測しながら、ユーザはエージェントの行動を監視しながらゴールや振舞いを修正していく。ユーザはエージェント系全体の振舞いに介入することができる。エージェントの行動の学習段階では有効となるインタラクションであるが、一般ユーザよりも設計者向きである。第3は自分の分身エージェントを置く方法である。ほかのエージェントの振舞いには直接関与せず、自分の分身エージェントにある程度の自律性を認めながら、各レベルで介入する。たとえばルーチ的な行動はエージェントに任せて、重要な判断を必要とする場面ではユーザが指示をだしたり、エージェントが決定をおくというような動作をする。

現在、上記の設計方針に基づいて、多人数ゲームを題材にモジュール間コミュニケーションのプロトコルなどを設計し、アーキテクチャの問題点の抽出をしている。A-HAの詳細については別の機会に詳しく報告する予定である [12]。

3.2 Script-based Behavior

Script-based なエージェントの振る舞いとしてはまず、Telescript[13] 的なものが考えられる。Telescript では、SCRIPT にしたがってエージェントがネットワーク上のリソースでプログラムを実行する。インタフェース・エージェントとしてはSCRIPT に従いながらインタラクションをいかにスムーズにできるかが問題となろう。

ヒューマン・マシン・インタラクション (HCI) を

スクリプトで規定することには、自由度のないインタラクションのなかでツールを使うことを想像し、反発があるかもしれない。しかしながら、いろいろなインタラクティブシステム（例えば[14][15]など）をみるとかなりユーザが演じるべき内容と流れが規定されている。スクリプトの中でユーザが制約をうけることは、むしろ場面ごとに適度な自由度をもつシステムの制御をすることになり、ユーザの習熟は早いと思われる。その典型的な例はビデオゲームである。また、インタフェース・エージェントのように状況を理解する認識系をもつシステムは、冗長な中で文脈依存の解釈をしなければならず、スクリプトベースのシステムにも有利な点が多いと考えられる[16]。

そこで、われわれはscript-basedなHCIをデザインする基盤となるスクリプトの記述方法とその際のイベント・マネージャを検討している[17]。とくに、マルチモーダル/ノンバーバルなインタラクションを行うようなシステムではセンサがスクリプトにしたがって効率よく動作し、かつ動作のタイミングについてゆるやかな制約をユーザに与えるようなイベントハンドリングが重要である。また、スクリプト制作をシステムのセンサやイベント記述と切り放すことによりシステム設計者でなくともスクリプトがかけられるようにすることができなければならない。これはセンサ類のツールキット化とスクリプティング・ツールのGUIをいかに作るかという問題まで含めて難題であるが、インタフェースの設計を一般まで開放することはHCI技術の進展において重要な課題である。スクリプトレイヤを独立させることはそれを加速できると考えられる。

Script-based エージェントは図1のpolicies/plans および Motivations のレイヤが Scripts レイヤに置き換わったものと考えて良い。ただしスクリプトの進行が支配的になるので、知識モニタ(knowledge monitor)の動作はイベントの制御が中心となる。

4 エージェントキャラクタのデザイン

4.1 キャラクタの動作のデザイン

インタフェース・エージェントの特徴である擬人性や生きいき性をもたせるにはCGキャラクタを作成して、エージェントが決めた行動に合せて動作させる必要がある。CGキャラクタは各種モデリングツールで比較的容易に生成できるようになったり、市販されるようになった。また、動きのデータもモーションキャプチャされたデータが使えるようになったが、エージェントの個性やタクスにあわせた

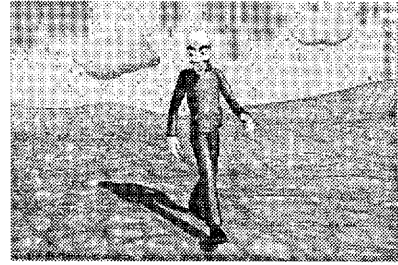


図3: 視線で足下に注意を促すように見えるキャラクタ (Believable Agent)

動きをもっと自由に簡単にデザインできるようにする必要がある。

我々は、キャラクタのアニメーションの制御・生成が、A-HAアーキテクチャと同じように本質的に階層的(図2)になっていると考える。ところが、それぞれの階層での微調整が他階層での制約を充足させるようにするのは非常に困難な作業である。例えば扉まで歩くようなアニメーションを作ったときに歩き方に個性をつけようと歩幅を変えた途端に、扉のところで足がそろわなくなってしまうようなことが生ずる。実際のCGアニメーションの制作現場ではこれらのステップが分業化され、アニメータの感性や密接なコミュニケーションや作業のやり直しなどで実現されている。階層ごとの制約を満足しながらパラメータの変更が他の階層に波及するシステムが必要である。

図4は歩行と走行のみを対象にして、そのような調整を簡単にしてお自然な動作を生成できるデザインツールである[18]。現在はインタラクティブなシステムになっているが、本ツールを用いてデザインした振る舞いの個性にしたがって、パラメータを順次変更するようなbehavior/actionモデルを作ることによって自然に歩行するキャラクタをインタフェース・エージェントとして登場させることができる。歩行、走行以外のアクションについてはこれからの検討課題である。

4.2 擬人化による認知性や身体性

エージェントに擬人性をもたせて、作業空間を可視化したり、身体性をあたえるとはどういうことだろうか。エージェントのタスクに依存するので具体例を示すのは困難であるが、いくつか例を示す。

4.2.1 エージェントによる作業空間の可視化

エージェントが有するあるいは作業している各種リソース(データ、知識ベース、対象物など)をバ

ターン化して表示することが効果的である。いわゆるビジュアライゼーションの技術が利用できる。また、エージェントの活動状況（煩雑度、困難度、結果の評価など）を表示することが有効と思われる。その際にインタフェースエージェントのジェスチャや表情で表すことができる対象と、作業空間を生成してその中でエージェントが作業しているところを示すほうが効果的な対象があると思われる。前者の例では、忙しく検索しているエージェントが走り回ったり、目がまわる表情をつけるなどがあるだろうし、後者の例では、検索対象のオブジェクトが空間を飛び回っているなかで、エージェントが関係する対象を取り出す動作を生成することができる。検索した結果はもらっても、いったいどんな母集団からとってきたのかわからないということはしばしば経験する。ここで、オブジェクトの量などが対象空間を大きさを直感的に示す手がかりとなることはいうまでもない。

4.2.2 エージェントによる身体性の付与

我々が物事を記憶したり考えたりするときに自分の体の全体や一部をリファレンスにすることがある。エージェントを擬人化することはその理解の補助になると考えられる。例えばガイドがが順路を示すときにいくつか方法がある。(i)文章で示したり、それを読み上げる、(ii)地図を示して現在地、目的地、順路を示す、(iii)位置関係を身ぶりですす、などである。このとき(iii)のように動作で示された順路は、自分の身体をリファレンスにするため記憶が容易であると考えられる。また、人間のジェスチャについては[7]などに示されるように、音声と共に(1)Iconic, (2)Metaphoric, (3)Beat, (4)Deictic (指示、ポインティング)に分類されるジェスチャが表出されることがわかっており、これらを情報がマルチモーダルなコミュニケーションを成立させていると考えられる。

4.2.3 リアリティ、ビリーバビリティ、擬人性、生きいき性?

believability や life-likeness を表現するポイントは何か。擬人化度をあげるためにリアルな顔や着衣やその動きの研究がおこなわれてきたが、リアルさを追求するほど現実とのギャップが目立つようになる。アニメーションに現れるマンガ的キャラクターでも十分にその世界では生きいきしていると感じることができるのはなぜであろうか。インタフェース・エージェントのキャラクターも現実をまねる必要はないが、believability や life-likeness のための要

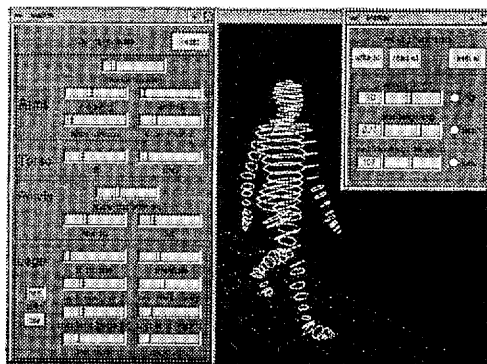


図 4: エージェントの歩行アニメーションデザインツールの画面

件を検討し、キャラクターにくみ込むことが必要であろう。

5 タスクとアプリケーション

インタフェースとして考えるときに、タスクを切り放して設計することは困難であるし、良いものを作ることはできないであろう。まえがきで述べたように、我々は発想支援や相互理解などについて具体的なタスクをあげ、そのときのインタフェース・エージェントがどのように振る舞うべきかを検討している。

メタ・ミュージアム[2]のガイドエージェントは拡張現実空間にあらわれて、ユーザの興味、スケジュール、位置、館内の状況を認識して次の展示への移動スケジュールをたててユーザへ提示する。その提示方法は、例えば、来館者をせき立てたり、鑑賞するのをじっと待ったり、あるいはユーザの足下への注意を促す(図3)などの動作を使うことが考えられる。

議論を活性化するエージェント[19]には、人間同士の会話に参加してタイミングをとって適切な意見を提示するという、もっぱら自律的で知的な活動を期待する。すなわち擬人化はあまり重要ではなくなる、とくに身体を全部表示する必要はなく、顔と手があれば充分であると考えている。むしろ、作業空間(エージェントが整理した議論空間)と状態(新しい意見を発する状態にあるかどうかの表象)を可視化すること、エージェントの見ている空間(world)がどこであることを明確にすることが重要である。

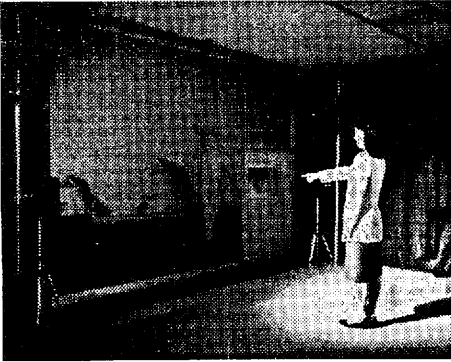


図 5: エージェントとのインタラクションの一例: メタミュージアムのサイバー世界をウォークスルーする際に現れるガイドエージェント(このキャラクターは自分の分身として位置と大きさのリファレンスの機能のみ有する単純なものである)

6 おわりに

例を示しながらインタフェース・エージェントについての基礎検討を行った。今後は、議論のなかで出てきた機能を実現する実験システムを作りながら考え方をリファインしていく予定である。

インタフェース・エージェントがもつべき機能とは、あるモデルやメタファに限定してインタフェースを提供するのではなく、いろいろなインタフェースを機能メニューとしてもって置いて、ユーザに応じて適応的に変えたり、カスタマイズできるようにすることができるようになることを、その自律性の設計に埋め込むことではないだろうか。とくに、使いこむにつれて、ユーザとエージェントの学習が進むにつれて、親密性を増せるような設計にしておくことが大事である。

ただの自動プログラムではなく、なぜエージェントと呼ぶかについては多くの議論があろう。我々は、エージェントとよぶことによって、知的なインタラクションが可能となるインタフェースや機能あるいはプログラム構成のデザインをどう考えたら良いかについてある指針を与えてくれ、具体的に考える場を提供してくれるのではないかと期待している。

謝辞

日頃ご指導いただく中津良平社長に感謝します。また本内容について日頃から議論していただく研究所ならびに第2研究室の皆様にも感謝します。

参考文献

- [1] 石田享: “エージェントを考える”, 人工知能学会誌, 10, 5, pp. 663-667(1995).
- [2] K. Mase, R. Kadobayashi and R. Nakatsu: “Metamuseum: A supportive augmented reality environment for knowledge sharing”, Intn'l Conf on Virtual Systems and Multimedia'96(1996).
- [3] 西本一志, 角康之, 間瀬健二: “新たな話題を提供し対話を活性化するエージェント”, 96年度信学会ソサイアティ大会(1996).
- [4] 角康之, 西本一志, 間瀬健二: “グループディスカッションにおける話題空間の可視化と発言エージェント”, 情報研報, 情報学基礎, F143-15, pp. 103-108(1996).
- [5] P. Maes: “Agents that reduce work and information overload”, Communications of the ACM, 37, 7, pp. 30-40(1994).
- [6] R. Nakatsu and K. Mase: “Tutorial Course on Life-like Believable Communication Agents”, SIGGRAPH96, New Orleans(1996).
- [7] J. C. et al.: “Animated conversation: Rule-based generation of facial expression, gesture & spoken intonation for multiple conversational agents”, SIGGRAPH94 Proceedings, Orlando, FL, pp. 413-420(1994).
- [8] K. Waters: “A muscle model for animating three-dimensional facial expression”, Computer Graphics, 21, 4, pp. 17-24(1987).
- [9] P. Maes, T. Darrell, B. Blumberg and A. Pentland: “The ALIVE system: Full-body interaction with autonomous agents.”, Proc. of the Computer Animation '95 Conference, Geneva, Switzerland(1995).
- [10] M. Turk: “Visual interaction with lifelike characters”, Int. Conf. on Automatic Face and Gesture Recognition, pp. 368-373(1996).
- [11] N. Tosa and R. Nakatsu: “For interactive virtual drama: Body communication actor.”, Proc. of 7th International Symposium on Electronic Art, Rotterdam, The Netherlands(1996).
- [12] S. F. et. al.: “Agent 99: Implementing a simple card game using agents”, 知能情報メディアシンポジウム(1996).
- [13] General Magic White Paper: “Telescript Technology”.
- [14] 間瀬健二: “動画像処理を用いた新しいマンマシンインタフェースの研究”, PhD thesis, 名古屋大学学位論文(1992).
- [15] A. P. Pentland: “Smart rooms”, Scientific American, pp. 54-62(1996).
- [16] 間瀬健二: “インタラクションのシナリオに基づくパターン認識理解 — エージェントメディアにおけるパターン認識・理解技術の役割”, 96年度信学会ソサイアティ大会(1996).
- [17] C. S. Pinhanez, K. Mase and A. Bobick: “Interval scripts: a design paradigm for story-based interactive systems”, Technical report, ATR MI&C tech. report(1996).
- [18] A. Bruderlin: “The creative process of animation human movement”, Knowledge-Based Systems(1996).
- [19] 西本一志, 角康之, 間瀬健二: “一参加者として対話に加わる対話活性化エージェント”, 信学技報 TL96-7 電子情報通信学会(1996).