

仮想都市環境におけるバーチャル・ヒューマンの開発

中西 英之^{†‡}, 板倉 豊和[†], 杉山 香織[†], 岡本 昌之[†], 山中 信敏[†], 石田 亨^{†‡}

[†]京都大学情報学研究科社会情報学専攻 [‡]JST CREST デジタルシティプロジェクト

1. はじめに

CGによって仮想的な人間を作るバーチャル・ヒューマン（以下，VH）の研究[1, 2]が盛んになってきている。我々は，都市空間をシミュレートする仮想環境である FreeWalk の実現に向けて[3, 4]，集団行動に参加し，人間との社会的インタラクションに従事することのできる社会的VHを開発中である[5, 6]。このVHは，これまで研究されてきた，ごく少人数の人間とのコミュニケーションに従事するエージェントや[1]，VH間のインタラクションに乏しいCGアニメーション[2]とは異なり，あらかじめ行動が予測できない多数のVHと，協調的に行動する能力が求められる。そのため，コンピュータが制御するVH（エージェント）の行動や対話を設計するのが困難になり，また，人間が操作するVH（アバター）の操作が難しくなる。

これらの問題は，これまで盛んに行われてきた，具現化対話エージェント[1]の研究や，仮想環境でのウォークスルーインタフェースの研究では積極的に取り扱われておらず，その解決には新たな手法が必要となる。我々の提案は，各VHの自律性を保ちつつも，VHの内側だけを設計するのではなく，外側にある仮想都市環境と合わせて一つのシステムとして設計しようというものである。

我々は日常生活の中で，人間の自然な行動パターンを，場面や状況とのセットで，膨大に記憶している。仮想都市環境は，日常生活の舞台である都市空間を再現したものである。そこで活動するVHの設計に，人間の行動パターンに関する知識を生かせるはずである。我々は，仮想都市環境内の場面，すなわち，そこにいるVHの集団と，それを取り囲む物体や建造物の属性やパラメータから，VHの自然な行動パターンを生成する手法を開発した。仮想空間は現

実空間と異なり，それらの属性やパラメータを，容易に取得可能である。

この手法は，多数のVHによる協調的な行動を再現するのに有効である。仮想空間を用いたチャットツールやオンラインゲームでは，各VHの行動があまり連携しておらず，その結果として，混沌とした場面が発生する。一方，我々の手法では，同じ場面が現実世界で起こったときの人間の行動に，VHの行動をできるだけ近づけることで，VHの集団の中に人間集団と同様の連携が生まれる。

また，この手法では，日常生活で遭遇する場面にもとづいてVHを設計する。そのため，この手法をさらに発展させ，VHの開発方法を，少数の開発者による仕組みの構築から，多数の利用者による振舞いの教育へと転換させることができるのではないかと考えている。

2. 都市の日常的場面にもとづく設計

我々は上記の手法をVHの，移動操作インタフェース，行動記述方式，対話戦略の設計に適用した。

人々が集団で行動するときは，他者との間合いを調節したり，列や円陣などの形に並んだりする。このような協調的な行動における，アバターの移動操作の煩わしさを軽減するために，ナビゲーション支援機能を開発した。アバターはリアルタイムで操作するのに対し，エージェントの行動は，事前にスクリプト言語で記述する。このスクリプトの実行時に，事前に予測できない仮想都市環境内の状況を反映するようにした。対話機能についても，これと同様に環境内の状況に従うようにした。特に，対話相手の非言語的合図を利用するよう実装した。

2.1 アバターのナビゲーション支援

この機能は，衝突回避といった物理的レベルの支援[7]だけでなく，同一集団内の他者との協調行動といった社会的レベルまで含むものである。社会的レベルのナビゲーション支援を実現する

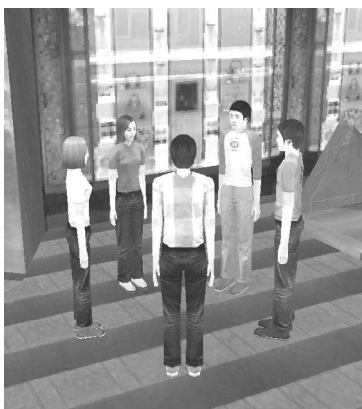
Designing Virtual Humans for Virtual City Environments
Hideyuki Nakanishi^{†‡}, Toyokazu Itakura[†], Kaori Sugiyama[†],
Masayuki Okamoto[†], Nobutoshi, Yamanaka[†], Toru Ishida^{†‡}
[†]Department of Social Informatics, Kyoto University
[‡]JST Digital City Project

ために、鳥や魚の群れをシミュレートする flocking (群れ) モデル[8]を参考に、フォーメーション (隊形) モデルを考案して使用した。

flockingモデルでは、各エージェントが、近傍にいる他のエージェントとの位置関係に関するいくつかの単純な規則に従って移動し、その結果として、群れとしての集団の隊形が浮かび上がってくる。フォーメーションモデルはこの逆で、各エージェントはまず、自分が所属している集団の隊形が何かを判別し、その中で自分がいるべき位置に向かって移動する。集団の隊形には、列に並んだり、出入口を通過したりするときの直線 (図1 (a)) や、会話をするときの円環 (図1 (b)) [9]などがある。このようにフォーメーションモデルは、人間が持つ状況認識能力と社会的習性を反映したものになっている。



(a) 直線状の隊形



(b) 円環状の隊形

図1. フォーメーションモデル

どの隊形かは、集団の形から判断する以外にも、各VHの向き、移動中かどうか、会話中かどうか、立っている場所、などからも判断できる。例えば同じ直線の形でも、横に並んで歩きながら会話している場合もあれば、建物の入口

の前で一列になって待っている場合もある。図1 (a)では、改札に向かって、全VHが一列になって歩いている。このように、隊形はいくつかの特徴によって分類でき、移動に関してだけでも、各VHの移動ベクトルの組み合わせによっていくつもの隊形に分類できる[10]。

隊形の特徴は同時に、その隊形を取っている集団に属するVHが満たすべき条件となる。同じ集団を形成しているVHはどれか、あるVHがどの集団に属しているのか、をその条件によって判断することができる。図1 (b)の円環状の隊形の場合は、円環の上で静止している、円環の中心方向を向いている、時々発話している、ことが条件となる。

この条件に従って、利用者が集団の一員としてアバターを移動操作しているのかどうかを判断し、一員であると判断した場合だけ、操作入力に手を加えるようにした。利用者によるアバターの移動操作を、支援機能が独断で変更することは、支援になる場合もあれば、妨害になる場合もあるため、このような機能が、不適切な介入を抑えるのに必要である。

2.2 エージェント制御コマンドの階層化

環境がそれほど変化しない場合、スクリプトに書く行動制御コマンドのパラメータは、最善と思われるもの一つだけに絞って選んでおけば良く、移動・発話・ジェスチャーなどに相当する原始的なコマンドを並べていくことで記述できる[11, 12]。ところが、FreeWalkのような常に状況が変化する環境では、場合分けの数が多くなるため、パラメータ指定の作業が繁雑になる。そこで、行動記述に使用するコマンドを抽象的なものにし、それを実行するとき、環境内の状況にもとづいてパラメータを設定した原始的なコマンドを生成・実行する仕組みにした。我々はこの仕組みを、「視線」のような単体のコマンドを生成するためではなく[13]、いくつもの種類の原始的コマンドを生成して連動させ、VHの振舞い全体を制御するのに用いた。

行動記述に使用する抽象的なコマンドの例として、他者に対して、周辺にある物について説明するコマンド (explain) (図2) を仮想都市環境 FreeWalk のエージェントに実装した。

FreeWalk のエージェントは、ビデオゲームのノンプレイヤーキャラクター (NPC) と同様に、有限状態機械をベースとしたスクリプト言語で制御できるようになっている[14]。現在は、ルールベースシステムと有限状態機械の組み合わせ

せであるシナリオ記述言語 Q が備わっている [15]. Q 言語ではスクリプトで使用するコマンドを自由に定義することができる. 我々は, 歩行 (walk) ・視線 (face) ・指差し (point) といった原始的コマンドに加えて, 説明 (explain) を追加した.



(a) 右にいる人に説明



(b) 左にいる人に説明

図2. explain コマンドの実行

説明コマンドの場合, 視線は説明する相手の方を向き, 腕は説明対象を差していなければならない. なおかつ, 自分の体で相手と対象の間を遮断してはならない. これらの制約を満たす姿勢と立ち位置を, 相手と対象と自分の位置関係から計算し, 歩行・視線・指差しコマンドのパラメータを生成する. 図2では, 説明する相手にとって案内板が見えやすくなるように, 相手の位置に合わせて自分の立つ位置を変えている. 自分の体の向きは, 相手の方を向くために回す首の間接角度の限界と, 案内板を指す腕の肩の間接角度の限界にもとづいて決める. 仮に, 同じことを原始的コマンドで直接記述しようとすると, このような計算処理を各記述者が, プログラミング言語より制限の多いスクリプト言語で書くことになり, 複雑な作業が発生する.

コマンドを抽象化せずに, 計算処理を各原始的コマンドの中で実行する中間的な解決策も考えられる. しかし, その場合は, 「説明」のためのコマンドであることを, 別途パラメータで

逐一指定する必要が生じ, また, スクリプト中のどの部分が「説明」という一つの行動に相当するのか把握しておかなければならない. 逆に, 抽象的なコマンドだけしかない場合は, スクリプト言語の柔軟性が損なわれる. よって, まず原始的なコマンドでエージェントを制御できるようにし, 後から必要な抽象的なコマンドを追加していくアプローチが良いと考える.

2.3 注視対象駆動型の対話シナリオ

FreeWalk のエージェントは, 人間 (アバター) と音声コミュニケーションを行えるように, 音声認識・合成エンジンを備えている. これにより, 任意の単語やフレーズを聞き取らせたり, 発話させたりするコマンドを用いて, Q 言語で対話シナリオを記述できる.

既存の対話エージェントの研究には, タスクを限定することで対話性能を向上させる手法が多い. 例えば案内タスクの場合は, 案内に必要な情報を集める対話に限定する [16]. 教育タスクの場合は, 教育用コンテンツの知識を教授する対話に限定する [17]. このような「タスク駆動型」の手法では, 対話相手から聞き取ったフレーズを解析し, キーワードマッチングの結果などにもとづいて返答を返す. このとき, フレーズの中に特徴的な単語が含まれているほど, 適切な返答を返すことができる.

ところが, 仮想都市環境が再現する街の中では, 会話のきっかけや, 話の文脈の背景が, 話者の頭の中ではなく, 周囲を取り囲んでいる環境中にあることが多く, タスク駆動型の手法が通用しない. そこで, 対話相手のアバターが注意を払っている対象を臨機応変に話題として取り上げるエージェントの対話シナリオを作成した. このエージェントは, アバターが今現在見ているものに注目する「注視対象駆動型」の対話を行う. この対話は, 相手の位置や視線方向を取得するコマンドを用いて, 注視対象をエージェントに推測させることで可能となる.

例えば, エージェントといっしょに歩いているアバターが「あれは何?」と発話したとき, この発話の文面をいくら解析しても適切な返答を返すことはできない. このとき, 我々のエージェントは質問の対象を環境中から見つけようとする. 図3では, アバターの視界の中心に映っている店について, エージェントが説明している.

従来のタスク駆動型対話では, それまでの対話の文脈にもとづいて返答を返す. そのため,

対話履歴から対話の流れをモデル化することで、エージェントの対話能力を構築する[18]。一方、我々の提案する、注視対象駆動型対話では、対話の文脈や流れは重要ではなく、その時その場の状況で対話相手が何を見ているのかが重要になる。そこで、仮想都市環境のどの場所で、誰と一っしょにいて、どの方向を見ているか、どのような返答を返す、といったルールを蓄積することで対話を可能にする。この蓄積作業を支援するためのオーサリング環境も開発中である。

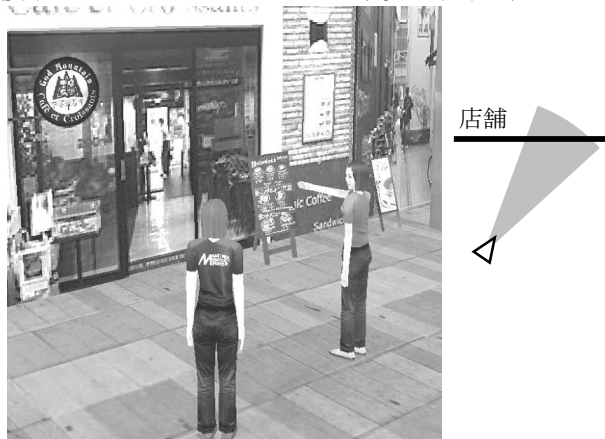


図3. 注視対象駆動型対話

3. おわりに

多数のバーチャル・ヒューマンが協調的に振舞う仕組みを、近接学、非言語的コミュニケーション、エスノメソドロジーなどの研究における人間集団の行動に関する知見を参考に開発した。アバター操作インタフェース、スクリプト実行系、対話シナリオ、などの各制御機構に、日常生活でよく見られる場面をルールとして埋め込むことで、ナビゲーションやコミュニケーションの能力を高めることができることを示した。今後はこのバーチャル・ヒューマンを、仮想都市空間での群集行動シミュレーションに用いる予定である。

謝辞 本研究は、JST CREST「デジタルシティのユニバーサルデザイン」プロジェクトの一部として進行中のものです。仮想都市環境の開発における、NTTサイバースペース研究所の筒口拳氏、京都大学の志水信哉氏、株式会社CRCソリューションズ、株式会社キャドセンター、株式会社数理システムの御協力に感謝します。

参考文献

[1] J. Cassell, J. Sullivan, S. Prevost, and E. Churchill Ed. Embodied Conversational Agents. MIT Press,

- 2000.
- [2] S.R. Musse and D. Thalmann. Hierarchical Model for Real Time Simulation of Virtual Human Crowds. IEEE Transactions on Visualization and Computer Graphics, Vol. 7, No. 2, pp. 152-164, 2001.
- [3] H. Nakanishi, C. Yoshida, T. Nishimura, and T. Ishida. FreeWalk: A 3D Virtual Space for Casual Meetings. IEEE MultiMedia, Vol. 6, No. 2, pp. 20-28, 1999.
- [4] T. Ishida. Digital City Kyoto: Social Information Infrastructure for Everyday Life. CACM, Vol. 45, No. 7, pp. 76-81, 2002.
- [5] K. Isbister, H. Nakanishi, T. Ishida, and C. Nass. Helper Agent: Designing an Assistant for Human-Human Interaction in a Virtual Meeting Space. CHI-2000, pp.57-64, 2000.
- [6] H. Nakanishi, S. Nakazawa, T. Ishida, K. Takanashi, and K. Isbister. Can Software Agents Influence Human Relations? - Balance Theory in Agent-mediated Communities -. AAMAS-2003, 2003 (to appear).
- [7] T. Kwon and Y. Choy. A New Navigation Method in 3D VE. VSMM-2000, A29T, 2000.
- [8] C.W. Reynolds. Flocks, Herds, and Schools: A Distributed Behavioral Model. SIGGRAPH-87, pp. 25-34, 1987.
- [9] A. Kendon. Spatial Organization in Social Encounters: the F-formation System. A. Kendon, Ed., Conducting Interaction: Patterns of Behavior in Focused Encounters, Cambridge University Press, pp. 209-237, 1990.
- [10] J. Allbeck, K. Kipper, C. Adams, W. Schuler, E. Zoubanova, N. Badler, M. Palmer, and A. Joshi. ACUMEN: Amplifying Control and Understanding of Multiple Entities. pp. 191-198, AAMAS-2002, 2002.
- [11] <http://www.vhml.org/>
- [12] S. Saeyor, H. Binda, and M. Ishizuka. Visual Authoring Tool for Presentation Agent Based on Multimodal Presentation Markup Language. IV-01, pp. 563-568, 2001.
- [13] S. Chopra-Khuller and N.I. Badler. Where To Look? Automating Attending Behaviors of Virtual Human Characters. Agents-99, pp. 16-23, 1999.
- [14] S. Cass. Mind Games. IEEE Spectrum, Vol. 39, No. 12, pp. 40-44, 2002.
- [15] T. Ishida. Q: A Scenario Description Language for Interactive Agents. IEEE Computer, Vol. 35, No. 11, pp. 42-47, 2002.
- [16] A. Stent, J. Dowding, J.M. Gawron, E.O. Bratt, and R. Moore. The Commandtalk Spoken Dialogue System. ACL-99, pp. 183-190, 1999.
- [17] J. Rickel and W.L. Johnson. Animated Agents for Procedural Training in Virtual Reality: Perception, Cognition, and Motor control. Applied Artificial Intelligence, 13:343-382, 1999.
- [18] E. Levin, R. Pieraccini, and W. Eckert. Using Markov Decision Process for Learning Dialogue Strategies. ICASSP-98, 1998.