

エージェントタイプに対応したバーチャル IO 形成システム

澤田志織^{1*}
Shiori Sawada¹

岨野太一¹
Taichi Sono¹

今井倫太¹
Michita Imai¹

¹ 慶應義塾大学

¹ Keio University

Abstract: ネットワークで繋がった複数のセンサからの情報を元に、エージェントが主観的に知覚したような知識表現を構成して与えるシステムを開発する。エージェントの体として扱うオブジェクト形状によるタイプ分類を行うことで、各タイプの身体の知覚範囲を考慮した上で汎用性を持った知識表現を実現する。このシステムにより、センサ処理の専門知識がなくともエージェントとのインタラクションシステムが製作できるよう支援する。

1 はじめに

今後、インタラクティブエージェントはより人々の生活に浸透していくことが予想される。PC など以外のデバイスもネットワークを介し情報を共有したりエージェントを搭載したりし始めていることから、ネットワークに接続されたエージェントの一般化はそう遠くないものと思われる。すると、技術者以外が自身で好みのエージェントとそのインタラクションを製作したいという需要も高まってくると考えられる。そこで本研究では、ネットワークに接続したエージェント（以下、ネットワーク型エージェント）を専門技術がなくとも簡単に製作できるようにする支援を目的としたシステムを開発する。

ネットワーク型エージェントと人間のインタラクションでは、エージェントが複数視点から情報を得ていると人間はエージェントの存在がわからず不安や混乱を感じてしまう。よって、エージェントに親しみやすくわかりやすいオブジェクトという形での身体が存在と、そのオブジェクトから想像される知覚と一致するような知覚を与えてやる必要がある。しかしこれを与える処理は知識や技術が要求されるため一般の人が全て行うのは困難であり、代わりに行うヴァーチャル IO (身体と知覚範囲) 生成システムが求められる。図 1 はヴァーチャル IO システムの有無による人間の感じ方の違いのイメージ図である。

ネットワーク型のエージェントに対し身体を与える先行研究は、デバイスにキャラクターをイメージさせるパーツを取り付ける手法 [1] や、エージェントがネットワーク上のデバイスの間を移動する様子を提示する手法 [2] がある。しかしこれらの研究は、ネットワー

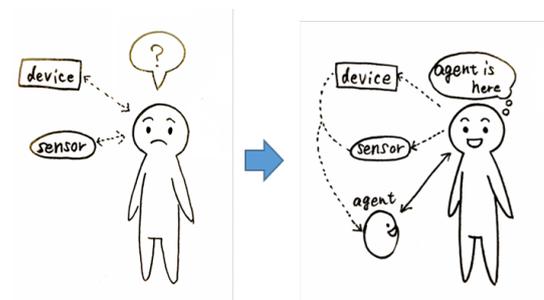


図 1: ネットワーク型エージェントとのインタラクションにおける問題と解決。

ク上の複数デバイスからの情報を同時かつ同列に扱う状況を想定しておらず、またエージェントの身体とするオブジェクトの一般性にも欠けるという問題がある。よって、より汎用的にネットワーク型のエージェントへ対応できるシステムが必要とされると思われる。

ネットワーク型エージェントにヴァーチャル IO を与える際にインタラクションの実現を目的とすれば、身体特性や知覚状態について詳細な表現や数値は必ずしも必要ない。

そのため本研究では、エージェント身体形状をタイプ分けし、ネットワーク上のセンサ情報を元にタイプの知覚に合致するような知識表現を構成し、エージェントに渡すヴァーチャル IO システムを開発する。これによってネットワーク型エージェントとそのインタラクション製作が容易になることが期待できる。また、ユーザが自由なオブジェクトをエージェントにすることが可能になる。

本稿では、セクション 2 で扱う問題にまつわる従来研究とその問題点、関連知識を紹介をし、セクション 3 で問題に対する本研究のアプローチを提案する。セクション 4 にて開発するシステムの概要と、システム

*連絡先: 慶應義塾大学理工学部情報工学科
〒223-0061 神奈川県横浜市港北区日吉 3-14-1
E-mail: sawada@ailab.ics.keio.ac.jp

が対応するエージェントタイプ、生成する知識表現について述べる。セクション5ではまとめと今後の展望を述べる。

2 関連研究と問題

2.1 ネットワーク型エージェント

2.1.1 本研究における定義

まず、ネットワーク型エージェントと非ネットワーク型エージェントの違いについて議論する。本研究ではその違いは、身体とセンサが明確かつ独立の物理デバイスに完結しているか否かである、と定義して扱うことにする。例えば、エージェントという用語から連想されやすい動物型や人型のロボットや多くの家電製品などは非ネットワーク型であり、単一の身体デバイスに装備されているセンサの視点によってしか外部の情報を得ていない。一方で、部屋内の複数デバイスが連携して動作するインテリジェントルームなどはネットワーク型であり、多数のセンサから得た情報が共有されており身体と知覚視点が複数あるといえる。

2.1.2 実装における問題

ネットワーク型エージェントと人間のインタラクションにおいて、周囲の複数デバイスにセンサがあり複数視点からエージェントが情報を得ている状況を想定すると、人間はエージェントがどこに存在してどこから知覚を行っているのか分からず、不安や混乱を感じるという問題がある。したがってネットワーク型エージェントとの自然なインタラクションを実現するには、人間が認識できる明確な身体を与えセンサ情報を身体に合わせて処理する必要がある。だがこの処理は専門知識や技術を要求するため、専門家以外がネットワーク型のエージェントを作るにはこの処理、いわばヴァーチャルIO形成処理、を行うシステムが必要と考えられる。

2.1.3 従来研究とその課題

ネットワーク型エージェントへインタラクションに適した身体を与えるシステムの先行研究はいくつかあるが、それらには汎用性の面で問題がある。デバイスにキャラクターをイメージさせるようなパーツを取り付ける手法 [1] では、ネットワーク中のデバイスのうち特定のデバイスを中心に扱っており、またその中心デバイスはパーツが取り付け可能な家電製品を対象としている。またエージェントがネットワーク上のデバイスの間を移動する様子を提示する手法 [2] では、一度に

動作するデバイスは一つに限っており、複数のデバイスを同時に扱ったりデバイス間で情報共有することは想定していない。

しかし、ネットワーク上のセンサやデバイスの使い分けに制約があったり、エージェントの身体オブジェクトの種類や外観が限定されたりするシステムでは、エージェントの使用目的やインタラクションの内容の幅も狭まってしまう。更にそのシステムを利用してエージェントを製作するには、制約への理解が求められ簡易な製作という目的から外れる。

2.2 知識表現

2.2.1 知識表現の必要性

単にヴァーチャルIOを生成し知覚を変更するだけでは、情報の内容はセンサで取得されるような数値のままであり、専門知識のない人がエージェントとのインタラクションを製作するのに充分わかりやすい表現とは言い難い。エージェント開発では、一般に知識表現と呼ばれる各種の情報の整理、表現方法が用いられており、それは本研究にも有効であると思われる。そこでどのような情報の内容、形式が用いられているのかについて紹介しその特徴を検討する。

2.2.2 知識の分類

知識工学では分野を問わず科学的知識の表現は、(1) 事物の知識 (物の定義や属するクラス、特性など)、(2) 事象の知識 (行為や出来事とその時系列、因果関係) (3) 方法の知識 (行為のための知識、規則など)、(4) メタ知識 (知識の性質や使い方)、の4つに分類できるとされている [3]。インタラクティブエージェントにおいては、エージェント自身の特性や扱うオブジェクトの定義などが (1) 事物の知識に相当すると考えられる。またセンサからのデータ情報とそこから推測できるエージェント、人物、環境の状況についての知識は (2) 事象の知識にあたる。エージェントが可能な行動内容や行動条件は、(3) 方法の知識であり、これら3種類の知識が適切に与えられることがエージェントの要件と考えられる。そして、知識を提供する人間が持つべき諸知識への理解が (4) メタ知識にあたるといえる。

2.2.3 知識表現の種類

知識表現の種類には、セマンティックネットワーク、フレーム、プロダクションルール、述語論理などがある [4] が、本研究におけるエージェント実装に用いるには問題があるものもある。実世界の汎用的な事象は述

表 1: フレーム表現の例.

human	
agent	is front

語論理のように真偽で表現できるものばかりではなく、またセマンティックネットワークは慣れていない一般の人には少々理解や扱いが難しい可能性がある。プロダクションルールで扱うような推論や行動条件はエージェント開発における方法の知識の表現には適しているであろう。フレーム表現は、例えば「人がエージェント正面にいる」という状況を表 1 のようにフレーム名 (human)、と属性 (agent) と属性値 (is front) によって表現する。このような Object-Attribute-Value で事実を表す形式は OAV 形式ともいう [5]。これは社会の多くの分野で用いられている情報整理の仕方と類似しており、一般ユーザにも十分理解しやすい表現と思われる。

3 提案

3.1 ネットワーク型エージェント

本研究では、より汎用性を持ったネットワーク型エージェント用ヴァーチャル IO 生成システムを開発することを目的とする。一般に、ネットワークの複数デバイスからの情報処理は、センサの配置や身体とするオブジェクトの形状特性を個別に検討する必要がある。しかし多くのインタラクションの実装においては、身体形状の特性はおおまかなタイプの分類で代用でき、また数値での詳細なセンサデータがなくとも人物や環境の状態にまつわる知識があれば十分である。

そのため本研究では、ネットワーク上のセンサ情報を元にエージェント身体の種類と知覚に合致するような知識表現を構成し、エージェントにヴァーチャル IO からの知覚情報として渡すシステムを提案する。これによって、ネットワーク型エージェントの製作は、インタラクションを中心にかつ身体オブジェクトからの連想に合った自然な記述で、容易にできるようになると言える。そして身体オブジェクトの選択とインタラクション内容の汎用性を実現する。

3.2 知識表現

セクション 2.2.2 で整理したエージェント開発における情報のうち、セクション 3.1 で提案したシステムでは、エージェントタイプの定義が (1) 事物の知識に相当すると考えられる。そしてセンサデータから生成する知識表現はエージェント、人物、環境の状況についての (2) 事象の知識にあたる。(3) 方法の知識は、インタラクション設計に際し製作者が作りたい内容に合わせエージェントに与える内容であるので、本システムが生成するのは適切でない。

セクション 2.2.3 で紹介したように理解しやすいフレーム表現を、本システムで表現方法に用いることとする。フレーム表現では表 1 と同じ内容をエージェントのフレームにおいて属性と属性値を変えて表現することもでき、着目する主体別に人間の認知に近い自由な内容の表現をすればより理解度が高まることが予想される。フレーム表現自体は単なるデータ構造であるため、その知識をどう扱ってエージェントが行動するかはエージェント側に一任する本システムには適していると考えられる。また OAV 形式で知識が与えられることにより、開発者は if ... then ... などのルール形

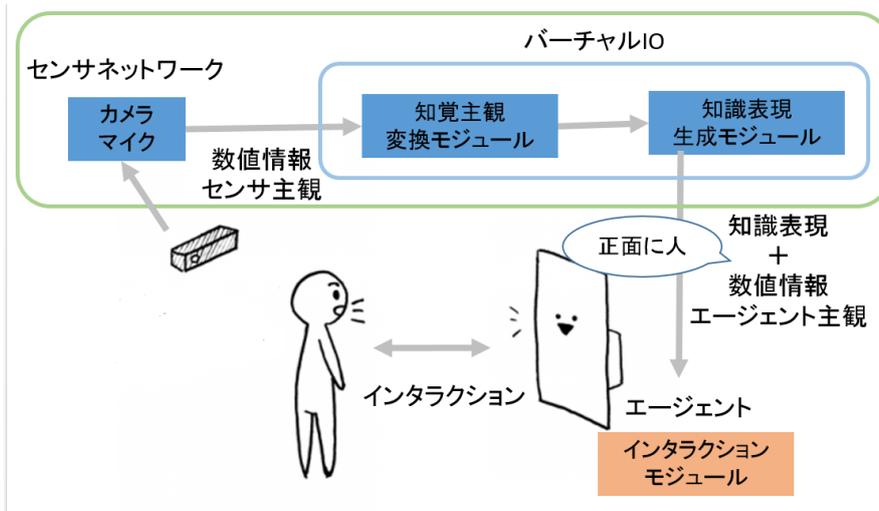


図 2: システム構成.

式でインタラクションの条件を記述する際には、条件部に OAV で記述される事実を用いれば記述も簡潔になることが期待できる。

4 システム構成

本研究で開発するシステムの概要は図2の通りであり、図2は平面イラストが身体のエージェントとインタラクションしている様子を例にとっている。

4.1 処理の流れ

センサネットワークから取得された元データは、センサ主観による数値情報である。これにまず、エージェントや人物の検出など必要な処理を行う。その手法の検討については本研究の主旨ではないため、マーカーによる認識や教師データを与えた検出など、各エージェントタイプに対し最も情報取得が容易なものを用いる。そして得られたデータをエージェントの主観によって知覚されたようにする数値変換を加える。その後、セクション3.3で述べる各エージェントタイプの知覚範囲に基づき、セクション3.4で具体例を挙げるような知識表現を生成する。そしてその知識表現と数値データをエージェント側のインタラクションモジュールに渡す。

4.2 セットアップ

本研究のシステム使用、実験において想定する状況は人間とエージェントが一对一でインタラクションする場合である。ネットワークセンサ入力装置として Microsoft Kinect V2 を1~4台と、首かけ型のマイクを用いる。Kinectはカラー画像、奥行き、人物データを取得できるため、複合的な情報を得るセンサネットワークを想定したシステムのセットアップに相当である。しかし Kinect は一定以上離れた音源からの音声取得能力には劣る。そこで理想的には環境内に設置されたマイクが望ましいが、本研究では便宜上ヘッドセット型のマイ



図3: 開発したシステムの使用状況例。

表2: エージェントタイプごとの知覚特性と行動条件。
○は必ず可能, △は条件あり, ×は常に不可能

	視覚	聴覚	移動回転	発話	自発移動
イラスト	△	○	○	△	×
箱	×	○	○	○	×
部屋	○	○	×	○	×

クを使用する。また出力装置としてスピーカーをエージェント身体オブジェクトの付近に配置する。

4.3 エージェントタイプ

本研究ではエージェントの身体オブジェクト形状の例として、以下の3つのタイプについて開発を行う。タイプごとに規定するのは、身体の形状(存在範囲, その指定方法), 知覚範囲(視覚, 聴覚, 移動と回転の体性感覚), エージェントの行動条件(発話, 自発移動)である。しかし他のタイプについても、これらを規定しそれに応じたモジュールに変更すれば当システムで対応可能である。以下で各タイプの身体形状特性と規定内容について述べ、表2に比較をまとめた。

4.3.1 平面のイラストタイプ

特別なオブジェクトを必要とせず、開発者が好みの画像を用意すれば良いため準備が簡単である。特に顔をモチーフにしたイラストであればコミュニケーションの対象として認識しやすいため、エージェントに与える身体のタイプに適している。

存在範囲の指定はキャリブレーション時にイラストの四隅の地点をユーザの手によって指定することによる。これによってエージェントの大きさの指定も兼ねる。その時点での画像を記録することによりトラッキングを行う。知覚範囲について、視覚は前方しか視認できず、人間の知覚と近い方が理解しやすいため図4

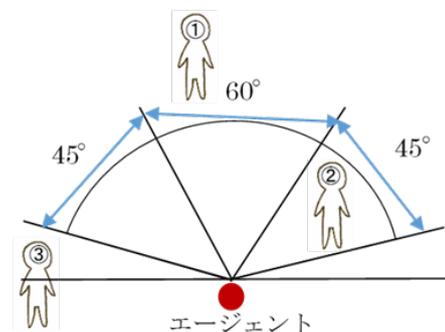


図4: 平面イラストのエージェントの視認範囲。

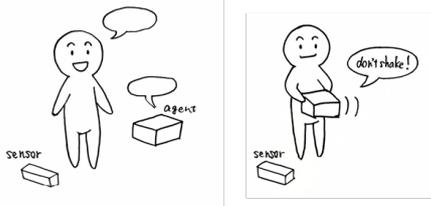


図 5: 箱のエージェントのイメージ図。
左: 箱型エージェントと話している様子
右: 人間がエージェントを動かしてインタラクションしている様子

のように前方 150°の角度の範囲を視野とすることにする。ただし目のパーツを人間が覆い隠した場合、視覚がなくなるものとする。聴覚については全方向から可能とし、音声取得は人間が装備するマイクによるので特に距離は制限しない。また体性感覚は人間がイラストを動かすことによる受動的な移動や回転をトラッキングにより知覚する。口のパーツが覆われていなければ発話可能とし、エージェント自身での移動はできないとする。

4.3.2 箱のエージェント

立体物の例として内部にエージェントがいる箱状のもの(図5)を扱う。キャラクター性を持たせる立体物としては、ロボットやぬいぐるみのように人間や動物等をモチーフとしたものが一般的であろうが、身体のパーツの有無形状などが多様化複雑化し、ユーザー側の認識も個人差が広がるため一般化が難しい。そこで本研究では、ユーザーがエージェントの知覚や能動性の詳細を想像せずインタラクションできるものとして、箱の中に入っている存在を例とする。

存在範囲は箱の角 8 箇所をキャリブレーション時に指定する。箱の色や形状を利用し移動回転を検知する。知覚は、視覚を持たず聴覚は全方位かつ距離無制限とし、体性感覚は受動的な移動と回転を知覚する。発話は制限せず、能動的な行動はできないとする。

4.3.3 部屋のエージェント

ネットワークで接続された複数センサがインタラクション時に利点を発揮するような例として、周囲の複数のセンサが広範囲を知覚しているようなインテリジェントルーム型のエージェントも扱う。身体であるオブジェクトとしては別途に物を用意するのではなく、部屋全体が身体といえる。

部屋全体であるため存在範囲の指定は不要だが、前提として部屋全体をセンサが把握している必要はある。知覚については、視覚と聴覚は部屋内の全てを知覚可

能である、一方移動回転は不可能のため体性感覚はない。発話は常に可能で、能動的な行動はなしとする。

4.4 知識表現の内容

本システムにおいてはフレーム形式(OAV形式)により知識表現を記述する。知識の主体となるフレーム名として、人物、エージェント、環境、の3つをエージェントタイプごとに設定する。例として、本研究でのイラストタイプエージェントにおける人物の場合は以下の属性と属性値を設ける。

- 位置方向: {不明, 正面, 右前, 左前}
- 距離: {不明, 近い, 遠い}
- 移動: {不明, あり, なし}
- 右手: {不明, 上がっている, 下がっている, 動かしている}
- 左手: {不明, 上がっている, 下がっている, 動かしている}
- 顔の向き: {不明, 正対, 右向き, 左向き, 後ろ向き}
- 発話: {不明, あり, なし}

例えば、人物の位置方向を表す属性値は図4のようにエージェントの視点を中心とする極座標と知覚範囲を元に決定し、図の人物に対し①「正面, 遠い」、②「右前, 近い」、③「不明, 不明」のような知識が生成される。ただし属性値の内、閾値にエージェントの状況を考慮した方が良いものはエージェント及び環境のフレーム属性の参照により変更を与える。上記の中では距離が「近い」、「遠い」はエージェントの大きさを考慮するものとする。また、センサ情報が取得できない場合はその旨を表現すべきであるため、全項目に「不明」の属性値を設ける。

5 おわりに

本研究では、ネットワークで繋がった複数のセンサからの情報を、エージェントが主観的に知覚したような知識表現に構成して与え、エージェント製作を容易にするシステムを提案した。その際、エージェントの身体として扱うオブジェクト形状によるエージェントタイプの分類を行うことで、各タイプの身体の知覚範囲を考慮した上で汎用性を持った知識表現が可能であると示した。開発したシステムで実装するのは静止物体の一例ではあるが、周囲のセンサを活用するロボットなどにも利用は可能である。

今後の計画としては、まずネットワーク型のエージェントであることを活かすようなインタラクションの内容を作成する必要がある。しかし、開発するシステムは開発者の負担軽減を目的とするものであるため、自身だけでインタラクションモジュールを実装しても効用を確かめるには不十分の可能性もある。そこで疑似コードを用いたコーディングを、本システムによる知識表現を使って行う場合と生のセンサデータを使って行う場合の2パターン被験者に行ってもらい、その差異を比較する実験による検証が可能であると思われる。また、知識表現の属性の項目をむやみに増やすなどしても、分かりにくくなり汎用性をかえて失う可能性もあるため、インタラクションの例を踏まえどのような表現が適当かを検討していく予定である。

6 参考文献

参考文献

- [1] Kentaro Ishii, Hiroataka Osawa, Seiji Yamada: An Authoring Tool for Interaction Method Evoking Character Imagery on an Object, *The 25th Annual Conference of the Japanese Society for Artificial Intelligence*, 2011
- [2] Mamoru Yamanouchi, Taichi Sono, Michita Imai: The Use of The BDI Model As Design Principle for A Migratable Agent, *The 4th International Conference on Human-Agent Interaction*, pp115–122, 2016
- [3] 吉村忠与志: 科学知識表現のための人工知能言語入門, 共立出版, 1991
- [4] 國藤 進, 羽山 徹彩, 中田 豊久: 知識基盤社会のための人工知能入門, 計測自動制御学会, 2012.5
- [5] 木下 哲男: 人工知能と知識処理, 昭晃堂, 2009.10