

人間・ロボット間のユビキタス・インタラクション環境

Development of Environments for Ubiquitous Human-robot Interaction

武田 英明 小林 展英 松原 慶幸 西田豊明
Hideaki Takeda Nobuhide Kobayashi Yoshiyuki Matsubara Toyoaki Nishida

奈良先端科学技術大学院大学 情報科学研究科

Graduate School of Information Science, Nara Institute of Science and Technology

Abstract: In this paper, we show three types of human-robot interaction which are important components to realize ubiquitous human-robot interaction. Firstly, we describe how intimate interaction is realized. We use gesture recognition, gesture generation, and speech generation for interaction. Secondly, we introduce a software agent called “watcher” to realize loose interaction. Watcher always looks at the environment to detect whether someone is requesting interaction. It also uses gesture recognition in a coarse level. Thirdly we provide mediators for cooperative interaction. A mediator is invoked each time interaction is requested in the environment. It can gather and enroll necessary agents for interaction by planning and sometimes by consulting other mediators.

1 はじめに

我々は人間と共存するロボット環境の構築について研究を行ってきた [4][3]。本研究ではこのなかで特に人間とロボットがどう関係すべきかという点について考察を行なう。

人間にとってはロボットが人間と同様な手段でコミュニケーションが図れることが望ましい。音声や身振りなどのいわゆるマルチモーダル・コミュニケーションは無論必要であるが、人間間のコミュニケーションは位置関係といった物理的な関係や特定・非特定といった論理的関係の変化にも対応可能な柔軟性がある。ここでは、このようなさまざまな物理的・論理的関係において可能な人間・ロボット間のインタラクションをユビキタス (遍在的)・インタラクション (ubiquitous human-robot interaction) と呼ぶことにする。

ここではユビキタス・インタラクション実現のために人間・ロボット間のインタラクションを3つの種類に分けて考える。まず第一は人間とロボットが直接向かい合っているインタラクションで、親密的インタラクション (intimate interaction) と呼ぶことにする。多くの研究での人間・ロボット間インタラクションはこの種類である。これに対して目の前にいないロボットとインタラクションを行ないたい場合がある。これを緩やかなインタラクション (loose interaction) と呼ぶことにする。さらにインタラクションを行なっていくうちにいまインタラクションに参加しているメンバー以外の協力が必要となる場合がある。これを協調的インタラクション (cooperative interaction) と呼ぶことにする。

以下ではこれらの3つの種類を統合的に実現する方法について述べる。その実現の鍵は環境の知能化である。すなわち、環境がロボットや人間に協力的にすることで、さまざまなインタラクションの実現を図る。

⁰ 武田 英明 奈良先端科学技術大学院大学情報科学研究科 〒630-01 生駒市高山町 8916-6 Tel (07437)2-5261 Fax (07437)2-5269 mail: takeda@is.aist-nara.ac.jp http://ai-www.aist-nara.ac.jp/

2 実世界エージェントのためのマルチエージェント・アーキテクチャ

我々はこれまで実世界エージェントのためのマルチエージェント・アーキテクチャを提唱してきた [4][3]。基本的な考え方は環境内の人間やロボット、機械などをエージェントとしてモデル化することで知識レベルのコミュニケーションを実現するものである。我々はこのコミュニケーションのために、対象・空間・行為に関するオントロジーを用意した。エージェント間コミュニケーションは KQML[1] で行なわれる。

ここでは主にソフトウェア・エージェントと実世界エージェントという二つの種類のエージェントを用意している。ソフトウェア・エージェントは実世界に対する直接のアクセス手段を持たないもので、エージェント間の協調などの役割を果たす。実世界エージェントはカメラやセンサ、モータやスピーカといった実世界の情報獲得、働きかけなどの手段をもったものである。

現在、実世界エージェントとしては6自由度のマニピュレータを持つ走行型ロボット Kappala、マニピュレータのない走行型ロボット Kappalb と1台の自動式棚、1台の自動式ドアを用意している。

3 親密的人間・ロボット間インタラクション

人間とロボットの直接のインタラクションである親密的インタラクションの手段としては、音声と身振りを用いた。身振りの理解は以下に示すような比較的単純な方法によって、人間の両手の動きのみを検知している。

まずエージェントは画像中の黒い部分を人間の頭の位置と仮定して、その両脇を手のある範囲として同定する (図1参照)。次にオブティカル・フローにより手の動作方向を検知する。検知結果は次のような両手の動作列である。

```
(move (object righthand)(direct up))  
(move (object lefthand)(direct down))  
(move (object bothhand)(direct left))
```

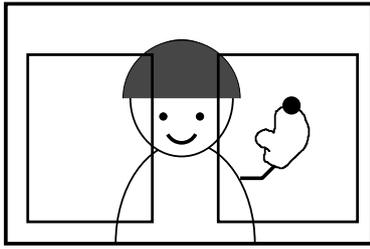


図 1: Extraction of motion area

```
(define wave"
  (struct
    (sequence
      (
        (move (object righthand)(direct right))
        (move (object righthand)(direct left))
        (move (object righthand)(direct right))
        (move (object righthand)(direct left))
      )
    ))
  (define shake
    (struct
      (sequence
        (
          (move (object righthand)(direct up))
          (move (object righthand)(direct down))
          (move (object righthand)(direct up))
          (move (object righthand)(direct down))
        )
      ))
    ))
```

図 2: Examples of definition of gestures

この検出した動作列を身振りの知識と比較することで、「握手する」、「手を振る」、「両手を動かす」などの身振りを同定する。身振りの知識は例えば図2のように記述される。

同定された身振りがどういう意味を持つかはそのときの状況に依存する。エージェントはインタラクション毎に定義された状況遷移ネットワークを用いることで、状況に依存した身振りの理解を行なう。エージェントは状況遷移ネットワークのいずれかのノードを現在ノードとする。状況遷移ネットワークの各リンクは人間の身振りと他のセンサの値に条件として付記してあり、それらが満たされた時、リンクの先のノードへ移動する。身振りは複数のリンクで利用可能であり、このための状況によって身振りの解釈が異なる。図3にこのインタラクションの例を示す。

実世界エージェントは主に2種類の行為を行なうことができる。一つは情報的行為 (informative action) で、環境に物理的変化を起こさない行為である。「はい」「いいえ」などの首を使った行為、「ばいばい」などの腕を使った行為、音声生成などが含まれる。もうひとつは影響を与える行為 (effective action) で、「ものを掴む」、「どこかへ移動する」などの行為である。これらのうちひとつのエージェントだけでできない行為は、第5章で述べるようにメディアータなどのソフトウェア・エージェントを用いて実現される。



図 3: An example of intimate interaction (“throw it out”)

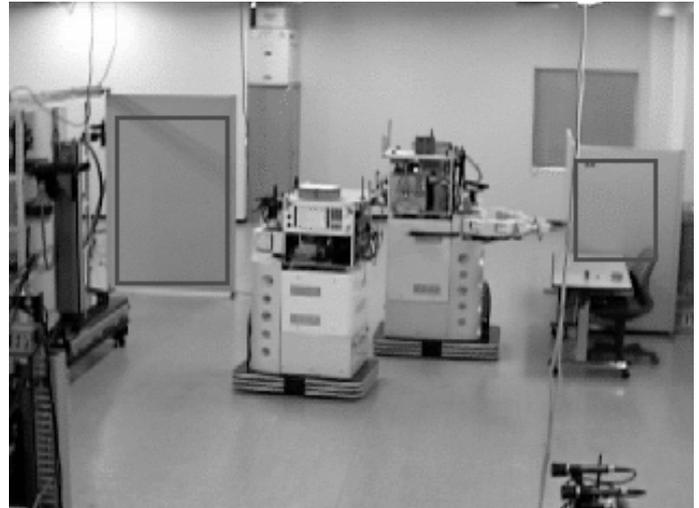


図 4: Scene by camera for watcher

4 緩やかな人間・ロボット間インタラクション

緩やかな人間・ロボット間インタラクションでは人間とロボットが離れている場合であり、ロボットは直接人間を見ることはできない。そこで、ここではウオッチャという部屋を常に監視するエージェントを導入する。ウオッチャはカメラを使い、部屋でおきる事象を監視する(図4参照)と共に、通信により他のエージェントからの通知を受ける。

ウオッチャは他からの要求を検知すると、タスクを生成してメディアータ(第5章参照)に渡す。要求の検知はカメラからの認識か他のエージェントからの通信による。カメラによる認識は親密なインタラクションと同様に人間の動作を認識する。現在、二つの領域(図4の矩形部)を監視している。認識された動作からタスクを生成するための知識の例を図5に示す。この例は「もし誰かが手を振っているならば、Kappalaがその人のところ

```
(define Come_on
  (content
    ((behavior wave)
     (source camera)
     (client ?human))
  )
  (task
    ((subject camera)
     (come (subject kappa1a)(destination ?human))
    )))
```

図 5: Knowledge on task composition



図 6: An example of loose interaction (a camera behind the robot detected human request and told the robot to go)

へいくというタスクを生成せよ。」ということを示す。

5 協調的人間・ロボット間インタラクション

インタラクションは状況に応じて、その参加者を増やす必要がある。例えば、ものを運べないロボットにものを運ぶよう人間が依頼した場合、ロボットは他のものを運べるロボットを呼び、そのロボットがものを人間にもってきて渡す必要がある。このインタラクションは2番目のロボットの導入することで完結する。

我々はこれまで開発してきたエージェント協調のためのエージェント・メディエータ [3] を拡張して、複数の非同期的要求を処理できるように拡張した。メディエータはエージェント同士の仲介をするものであるが [2]、ここではあるエージェントが生成したタスクを補完・分解してエージェントが実行可能にする役割を果たしている。

本研究ではメディエータはインタラクション毎に生成され、各メディエータはそのインタラクションが必要なエージェントを確保し、制御しようとする。もし、メディエータの要求同士が衝突する場合はメディエータ同士が通信して解決する。インタラクションに対する要求は以下のように処理される。

1. 与えられた要求を実現するタスクを生成する。タスクは行為として表現される。ただし行為のすべての属性が埋められている必要がない。タスクは

```
(fetch (object solaris)
       (destination human1))
      ↓
((move (subject Kappa1b)
       (from (at Kappa1b))
       (to (in_front_of Rack)))
 (handover (object solaris)
            (from_place (on Rack))
            (to_place (on Kappa1b))
            (subject Rack))
 (move (subject Kappa1b)
       (from (at Kappa1b))
       (to (in_front_of Human2)))
 (tell (subject Kappa1b)
        (at Kappa1b)
        (content (talk Here_you_are))))
```

図 7: An example of completion and decomposition of tasks

ウオッチャが検知した場合ウオッチャが、それ以外はその要求を出したエージェントが生成する。

2. ウオッチャはタスク毎にメディエータを生成してそのタスクを渡す。
3. 対象・空間・行為に知識を利用してタスクを補完・分解する。これはメディエータのプランナ部で行なわれる。この結果、図 7 に示すような行為列がプランとして生成される。あるメディエータが必要とするエージェントがすでに他のメディエータによって確保されている場合、メディエータは他のメディエータにどれだけそのエージェントを確保するかを問い合わせ、その長さによって待つかそのエージェントを使わないかを決定する。
4. プランを実行する。これはメディエータのエクゼキュータ部で行なわれる。行為列の行為を順に1つづつ実行するよう依頼し、その終了をまって次を実行する。プランに関わるすべてのエージェントがプラン開始時からそのエージェントの行為の終了まで確保される。なんらかの事情でプラン中の行為が失敗に終ると、プランナ部に失敗が伝えられ、再プランニングが行なわれる。

図 8 は本節の冒頭で述べた協調的インタラクションがどのように実行されるかを示す。図中の番号は以下の番号に対応する。

1. 人間が行動型エージェント Kappa1a に身振りでもニユアルを取るよう依頼する。
2. 行動型エージェントはその要求を理解するが、そのマニュアルは棚にあり自分は棚から受渡しができないので、自分では解決できないことがわかる。そこで、タスクを生成してウオッチャに渡す。
3. ウオッチャはメディエータを一つ起動してそのタスクを渡す。
4. メディエータはタスクを補完・分解して、Kappa1b がマニュアルを取りにいくというプランを生成し、実行する。

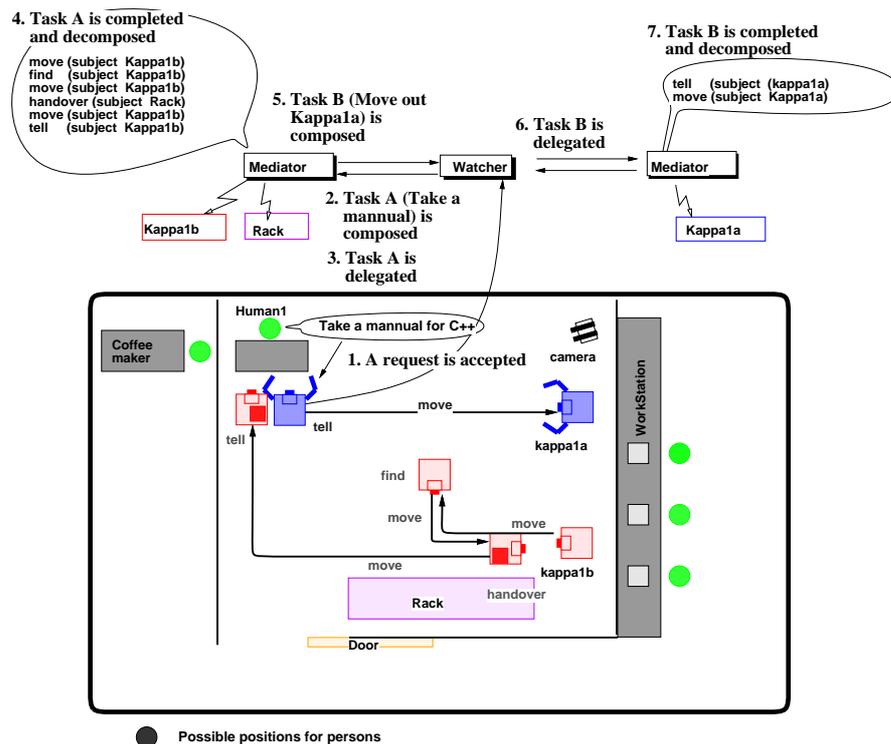


図 8: An example of cooperative interaction

5. メディエータは先のプランに実行には Kappa1a が邪魔であることがわかる。そこで Kappa1a を移動するタスクを生成してウオッチャに渡す。
6. ウオッチャは新たなメディエータを生成してタスクを渡す。
7. 2 番めのメディエータもタスクを補完・分解して、プランを生成して、実行する。

協調的なインタラクションは単に複数のロボットが利用可能という意味だけでなく、ロボットの物理的能力の限界を補完するという意味でも重要である。すなわち、自分でできないことも他のエージェントに依頼して結果的に解決可能であるので、人間はロボットの能力を推測せずに単に目の前のロボットに依頼するだけでよい。

6 おわりに

ここでは遍在的な人間・ロボット間のインタラクションを実現するために3つのインタラクションの種類を挙げると共にそれらを統合的に実現する枠組について述べた。その実現の鍵は実世界をマルチエージェントとしてモデル化することにある。

その利点のひとつは、実世界内のインタラクションだけでは難しい点をエージェント間通信に担わせるということである。たとえば、部屋を常に観察するということはロボット単体では難しいタスクであるが、それを専門に行なうエージェントとそれとの通信によって容易に実現可能である。

もう一つの利点は、ソフトウェアエージェントとも含めて構成することにより物理的構成とは異なる機能分散がはかれるということである。多くの研究ではロボット

をどれだけ自律的にできるかが研究課題とされてきた。本研究ではロボットそれ自身の知能化(ここでの実世界エージェント)の程度は必ずしも重要ではない。むしろ、環境の知能化、仲介サービスなどのソフトウェア・エージェントなどと一緒に働く時の振舞いの高度化が重要である。すなわち、知能的振舞いを実現する機能をエージェントに分散し、それらの協調として実現することができる。

参考文献

- [1] Tim Finin, Don McKay, Rich Fritzson, and Robin McEntire. KQML: An information and knowledge exchange protocol. In Kazuhiro Fuchi and Toshio Yokoi, editors, *Knowledge Building and Knowledge Sharing*. Ohmsha and IOS Press, 1994.
- [2] Hideaki Takeda, Kenji Iino, and Toyooki Nishida. Agent organization and communication with multiple ontologies. *International Journal of Cooperative Information Systems*, Vol. 4, No. 4, pp. 321–337, December 1995.
- [3] Hideaki Takeda, Koji Iwata, Motoyuki Takaai, Atsushi Sawada, and Toyooki Nishida. An ontology-based cooperative environment for real-world agents. In *Proceedings of Second International Conference on Multiagent Systems*, pp. 353–360, 1996.
- [4] 武田英明, 岩田浩司, 鷹合基行, 沢田篤史, 西田豊明. 共有知識に基づく実世界エージェントの協調機構. 人工知能学会全国大会(第10回)論文集, pp. 453–456, 1996.