日本語 Wikipedia オントロジーと動作オントロジーを利用した人型ロボットの対話と動作の融合

Intelligent Humanoid Robot with Japanese Wikipedia Ontology and Robot Action Ontology

小林昭太郎　後藤あいり　玉川奨　森田武史　山口高平

Shotaro KOBAYASHI, Airi GOTO, Susumu TAMAGAWA, Takeshi MORITA and Takahira YAMAGUCHI

慶應義塾大学
Keio University

Abstract: WioNA (Wikipedia Ontology NAo) is proposed to build much better HRI by integrating four elements: Japanese speech interface, semantic interpretation, Japanese Wikipedia Ontology and Robot Action Ontology. By giving Japanese Wikipedia Ontology to humanoid robot named "Nao" as wisdom, Nao can dialogue with users on many topics of various fields. In addition, aligning Robot Action Ontology for robot action control with Japanese Wikipedia Ontology enables Nao to perform related actions to dialogue topics. Case studies show us how HRI goes well in WioNA with these topics.

1. はじめに

現在，日本では産業用ロボット市場は頭打ちになっており，それに伴い，サービスロボットに注目が集まっている．これまで日本でもさまざまなサービスロボットが開発され，製品化されている．そして，少子高齢化などを考慮すると，今後も我が国でサービスロボットに対する需要は必ず高まっていくことは明らかある．しかし，人間とのインタラクション（Human Robot Interaction）が満足にいかないという問題がある．私たちは，この最大の原因は現状のサービスロボットは知識を持たないためであると考え，オントロジーの視点からアプローチする．

近年，Web 上の百科事典 Wikipedia が，豊富な語彙，情報の即時性により，新たな情報資源として注目されている．しかし，Wikipedia は人間が読むために記述された情報資源であり，ロボットは Wikipedia 自体を理解することはできない．このような事実を踏まえ，本稿では，セマンティック Web の最も重要な要素の一つであるオントロジーに着目した．オントロジーは，クラス・インスタンス・プロパティの３要素から構成され，コンピュータに意味的な理解を可能にする．さらに，高い表現能力，スケーラビリティ，高度な推論メカニズムを有する．したがって，日本語 Wikipedia から大規模オントロジーを構築しロボットに与えることによって，Wikipedia の情報を利用して人間と対話を行う，すなわちロボットが人間のインテリジェンスな振る舞いを模倣することが可能になる．さらに，ロボットの実行可能動作を体系的にまとめ構築した動作オントロジーを，この日本語 Wikipedia から構築した大規模オントロジー（以下，日本語 Wikipedia オントロジー）とアライメントを取ることによって，対話と動作の融合を実現することができる．ロボットの分野でオントロジーを使った研究はまだまだ少ないが，これらが本稿でオントロジーに着目した理由である．

本稿では，フランスの Aldebaran robotics 社[*1] 製人型 Nao を利用した，日本語 Wikipedia の情報に基づく対話と動作の融合を実現するシステム WioNA（Wikipedia ontology NAo）について述べる．

2. システム概要

WioNA は，日本語音声インタフェース，意味理解，日本語 Wikipedia オントロジー，動作オントロジーの４要素から構成される．日本語音声インタフェースは，オープンソースソフトウェアを組み合わせ，現状日本語能力を全く持たない Nao のために独自に開発した．意味理解については，オントロジーとの親和性が高い意味文法を採用した．日本語 Wikipedia オントロジーは，日本語 Wikipedia から半自動的に

連絡先：小林昭太郎
慶應義塾大学理工学部管理工学科
〒223-8522　神奈川県横浜市港北区日吉 3-14-1
TEL:045-566-1614
E-mail： s_kobayashi@ae.keio.ac.jp

*1 http://www.aldebaran-robotics.com/en/

構築し，意味理解のための辞書として利用した．これにより，Nao は，日本語 Wikipedia の膨大な情報に基づいて人間と対話することが可能になる．さらに，本稿では，ロボットの動作制御，および動作生成のために実行可能動作を体系化した動作オントロジーを，日本語 Wikipedia オントロジーとアライメントを取ることによって，対話と動作の融合を実現した．WioNA の概要を図 1 に示す．

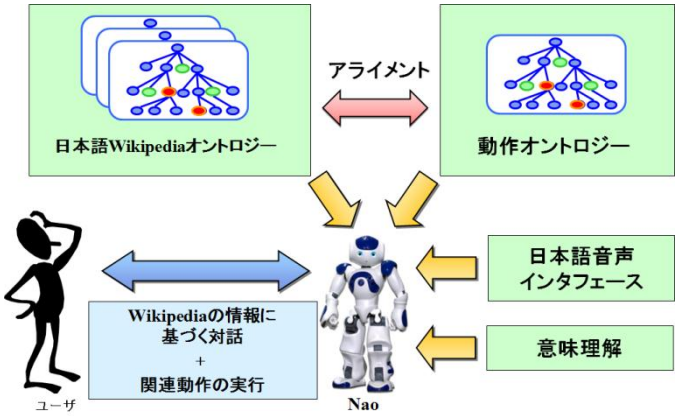

図 1 WioNA 概要

3. 人型ロボット Nao

Nao は，フランスの Aldebaran robotics 社によって開発された人型ロボットであり，世界中の研究機関で利用されている．Nao の最大の特徴は，Linux を搭載しフルプログラマブルであるという点である．これが本研究で Nao を利用した最大の理由である．また，マイク，スピーカー，LED，センサなど多くのデバイスを搭載する．さらに自由度は 25 あり，非常に複雑な動作を実行でき，その動作は付属ソフトウェアの Choregraphe を使うことにより，GUI によって容易にデザインすることができる．この点についても本研究で Nao を利用した大きな理由である．

4. 関連研究

本章では，オントロジーをロボティックス分野の研究に利用している研究について述べる．Johnston[1]らは，オントロジーをシンボルグラウンディングに利用し，OBOC（Ontology Based Ontology Categorization）というシステムを開発し，それをロボカップで評価している．OMRKF（Ontology-based Multi-layered Robot Knowledge Framework）[2]は，ロボットのための知識フレームワークであり，ロボットのためのオントロジー，人間のためのオントロジー，および多くのルールから構成される．Suh らは，これをロボットの運搬タスクを通して有用性を検証している．

5. 日本語音声インタフェース

Nao は，現時点では，日本語音声インタフェースを持たない．(今後のアップデートにより追加される予定）そのため，オープンソースソフトウェアを組み合わせ，音声合成・音声認識モジュールから構成される独自の日本語音声インタフェースを構築した．

5.1 音声合成モジュール

音声合成モジュールの開発には，Gtalk[*2] と SoX[*3]（Sound eXchange）の２つのソフトウェアを利用した．Gtalk は，Galatea Project[*4][3]により開発された日本語の漢字かな混じり文から音声を合成できるソフトウェアである．一方，SoX は，音声ファイルを変換するコマンドラインツールである．

音声合成モジュールの構成は，以下の通りである．まず，漢字かな混じり文から Gtalk と SoX により，Wav ファイルの形式に変換する．生成された Wav ファイルを Nao の Linux に，プログラムから，無線ネットワーク経由でアップロードする．それを Nao のスピーカーから再生することにより，Nao が話す仕組みである．しかし，日本語テキストを直接 Gtalk で合成した場合，文字数が多いと合成に時間がかかり，対話に大きな支障をきたす．したがって，テキストを句点（。）で分割し，文単位で合成処理を行うことにより，wav ファイルの合成と再生を並列して処理する．これにより，開発した日本語合成モジュールでは，テキストの長さに関わらず，Nao が約 1 秒程度で話し始めることが可能になった．

5.2 音声認識モジュール

音声認識モジュールは，Julius[*5] と Adintool[*5] という２つのソフトウェアを使い，開発した．Julius は，数万語を対象としたほぼリアルタイムな大語彙連続音声認識を行うことのできる認識エンジンである．Adintool は，Julius 付属のソフトウェアの 1 つで，音声ファイルの記録，分割，送信，受信を行うためのツールである．

音声認識モジュールの構成は，以下の通りである．まずユーザが Nao に話しかけると，Nao に搭載されているマイクが感知したデータを Adintool により，Julius が動く PC に無線ネットワークを通して送信する．それを Julius が受信し解析することにより，認識結果が得られる．なお，Nao の Linux 上で動作する Adintool は，PC 上で動く Python プログラムから paramiko[*6] を用い SSH 経由で実行した．

*2 http://sourceforge.jp/projects/galateatalk/

*3 http://sox.sourceforge.net/

*4 http://hil.t.u-tokyo.ac.jp/~galatea/index.html

*5 http://julius.sourceforge.jp/

*6 http://www.lag.net/paramiko/

5.3 意味理解

現在，意味理解の手法としてはさまざまな手法が提案されているが，本稿ではその中で意味文法を利用した．意味文法は，最新の技術ではないが，オントロジーとの親和性が非常に高く，また実装が容易であるというメリットがある．意味文法は，事前に文法パターンと単語辞書を定義する必要があるため，一般的には汎用的なトピックにおける対話には利用できない．本稿では，ユーザが話す文法パターンを限定し，かつ単語辞書を対話のフェーズに合わせてオントロジーを参照し動的に生成することにより，これを可能にした．単語辞書の動的生成については，自然言語処理ライブラリ MeCab[*7] を利用した．

6. 日本語 Wikipedia オントロジー

近年，その豊富な語彙，即時更新性により，新たな情報資源として Wikipedia が注目されている．本稿では，日本語 Wikipedia から機械可読な大規模オントロジーを生成し，Nao に意味理解用の辞書として与える．これにより，ユーザと Wikipedia の膨大な情報に基づいた対話を行うことが可能になる．Wikipedia は，本文など構造化されている部分と，Infobox など構造化されていない部分が共存する半構造化資源である．したがって，オントロジー生成のためのコストが小さい．そのため，DBPedia[4]や YAGO[5]など，Wikipedia からオントロジーを学習する研究が数多く行われている．しかし，これらの研究は，英語 Wikipedia を情報ソースとして利用しているため，日本語での対話を前提としている WioNA で利用することは困難である．このような理由で，本稿では，日本語 Wikipedia オントロジーを構築した．オントロジーは，クラス，インスタンス，プロパティの３要素から構成されるが，これらは，URI（Uniform Resource Identifire）によって，一意に識別される．日本語 Wikipedia オントロジーでは，それぞれ wikionto_class[*8]，wikionto_instance[*9]，wikionto_property[*10] という接頭辞を用いた．日本語 Wikipedia オントロジーは，以下の５つの要素により構成される．

1. Is-a relationship (rdfs:subClassOf)
2. Class instance relationship (rdf:type)
3. Infobox triple (owl:Object/DatatypeProperty)
4. Property domain (rdfs:domain)
5. Synonyms (skos:altLabel)

日本語 Wikipedia オントロジーの学習方法については，本稿の主題ではないため，WioNA で主に利用している【Is-a 関係】，【クラス-インスタンス関係】，【Infobox トリプル】について簡潔に示す．詳細については，参考文献[6]に示されている．

Is-a 関係（rdfs:subClassOf）

日本語 Wikipedia は，記事の分類を目的とした階層構造を持つ．しかし，下位カテゴリと上位カテゴリとの関係は，性質の継承という観点から捉えた Is-a 関係ではない場合が多いため，洗練が必要である．そのため，カテゴリ階層に対する文字列照合，カテゴリ名と Infobox テンプレートの照合，目次見出しのスクレイピングの３つの手法を用いて，Is-a 関係を構築した．

クラス-インスタンス関係

日本語 Wikipedia の一覧記事には，関連する事物が列挙されている．したがって，記事名をクラス，列挙された事物をインスタンスとみなしスクレイピングを行うことで，クラス-インスタンス関係を抽出することができる．この手法により，421,989 個のクラス-インスタンス関係を抽出した．

Infobox トリプル

Wikipedia の Infobox における「記事－項目－値」という三つ組は，オントロジーにおける「インスタンス－プロパティ－プロパティ値」という三つ組と捉えることができる．この方法により，1,203,404 個のトリプルを構築した．

図 2 に日本語 Wikipedia オントロジーの一部を示す．図 2 では，クラス，インスタンス，プロパティは，それぞれ青色，黄色，緑色で示し，接頭辞は省略した．また，図 2 を始め本稿における全てのグラフは，RDF モデル[*11] に準拠している．

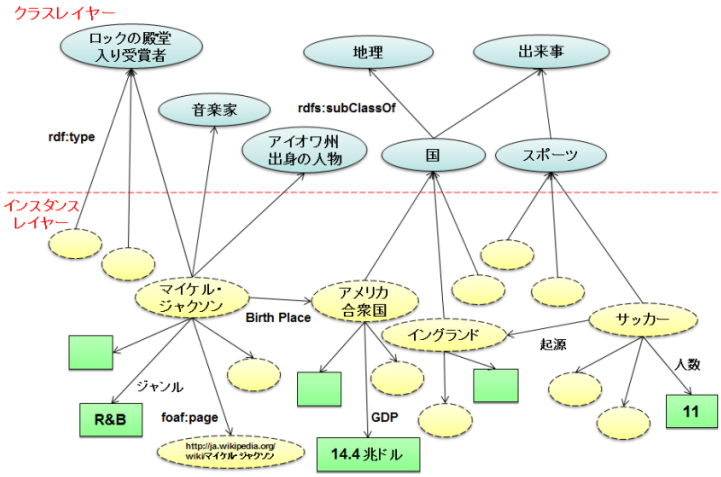

図 2　日本語 Wikipedia オントロジー（抜粋）

*7 http://mecab.sourceforge.net/

*8 wikionto_class: http://www.yamaguti.comp.ae.keio.ac.jp/wikipedia_ontology/class/

*9 wikionto_instance: http://www.yamaguti.comp.ae.keio.ac.jp/wikipedia_ontology/instance/

*10 wikionto_property: http://www.yamaguti.comp.ae.keio.ac.jp/wikipedia_ontology/property/

*11 http://www.w3.org/TR/rdf-concepts/

7. 動作オントロジー

動作オントロジーについては，Nao の実行可能動作を体系的にまとめ手動で構築した．動作オントロジーは，Nao の動作制御，および動的な動作生成に利用した．動作オントロジーでは，クラス，インスタンス，プロパティは，それぞれ roboonto_class[*12]，roboonto_instance[*13]，roboonto_property[*14] という接頭辞を用いて記述した．（以下略）

7.1 Is-a 関係，クラス-インスタンス関係

図 3 に，動作オントロジーの Is-a 関係，クラス-インスタンス関係の一部を示す．

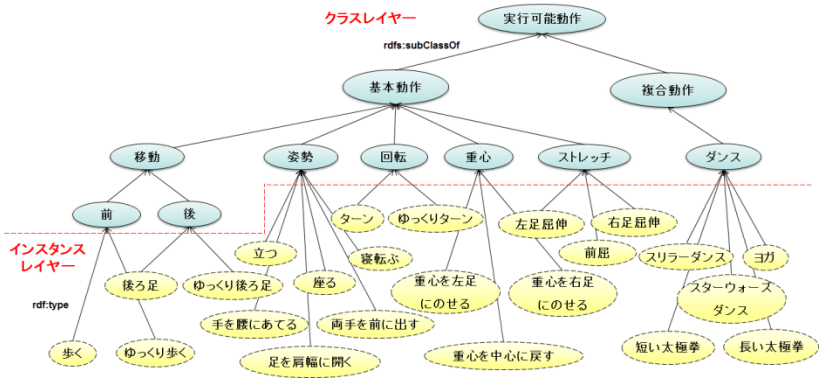

図 3　動作オントロジー（抜粋）

図 3 の黄色の部分がインスタンスであり，Nao の動作実行ソフトウェアモジュールが対応付けられている．一方，青色の部分は複数の実行可能動作をまとめるクラスを表す．また，図 3 から分かるように，全てのインスタンスは，基本動作クラスと複合動作クラスのいずれかに属する．基本動作クラスのインスタンス（以下，基本動作）は，単一の動作ファイルによって実行できる動作，それ対して，複合動作クラスのインスタンス（以下，複合動作）は，複数の基本動作を連続して実行することによって，実行できる動作である．このように実行方法に基づいて分類することにより，実装が容易になる．

7.2 インスタンストリプル

動作オントロジーのインスタンスに，動作に関するデータをトリプルとして記述した．これについて，基本動作と複合動作に分けて示す．

7.2.1　基本動作

基本動作は，Difficulty，Length，Path という 3 つのデータタイププロパティ（値がリテラルであるプロパティ）を持つ．Difficulty，Length プロパティは，それぞれ各動作の難易度，所要時間を表す．また，基本動作は単一の動作ファイルによって実行されるため，その動作ファイルのパスを Path プロパティで示した．さらに，基本動作は，他の基本動作に PreCondition というオブジェクトプロパティ（値がオブジェクトであるプロパティ）でリンクしており，インスタンスネットワークを形成する．具体的に述べると，ある基本動作を実行するための前提条件を，PreCondition プロパティ，および順序を問わないコンテナモデルの rdf:Bag で記述した．この前提条件に基づき，複合動作における短縮・延長バージョンを動的に生成できる．図 4 に【重心を中心に戻す】という基本動作におけるインスタンストリプルを示す．

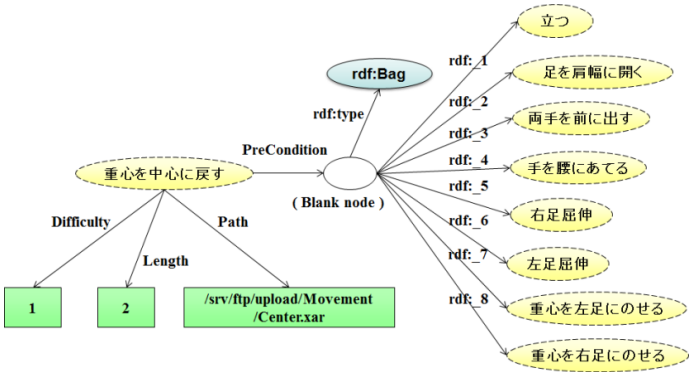

図 4　基本動作におけるインスタンストリプル

7.2.2　複合動作

複合動作はいくつかの基本動作を連続して行うことにより実行される．そのため，各複合動作のアクションフローを Action_Flow というオブジェクトプロパティ，および順序を問うコンテナモデルの rdf:Seq を用いてモデル化した．すなわち，複合動作はその構成要素である基本動作にリンクしており，インスタンスネットワークを形成している．図 5 に，【短い太極拳】という複合動作のモデルを示す．

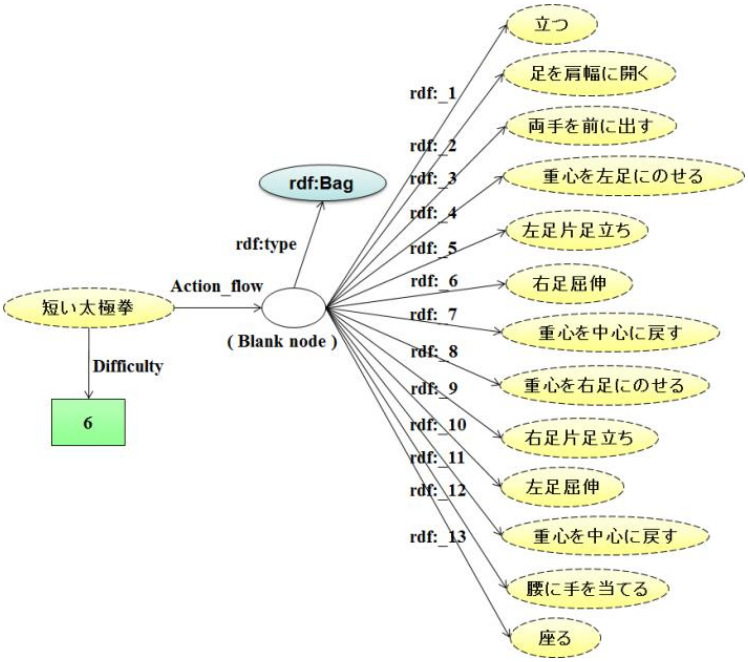

図 5　複合動作におけるインスタンストリプル

---

*12　roboonto_class: http://www.yamaguti.comp.ae.keio.ac.jp/robot_action_ontology/class/

*13　roboonto_instance: http://www.yamaguti.comp.ae.keio.ac.jp/robot_action_ontology/instance/

*14　roboonto_property: http://www.yamaguti.comp.ae.keio.ac.jp/robot_action_ontology/property/

7.3 アライメント手法

動作オントロジーを日本語 Wikipedia オントロジーとアライメントをとることにより，対話だけでなく，関連動作を抽出し実行することが可能になる．すなわち，動作オントロジーと日本語 Wikipedia オントロジー間のアライメントは，動作と対話をつなぐ役割を担う．理想的には，これらのアライメントを計算機により自動的に処理するのが望ましい．しかし，動作オントロジーと日本語 Wikipedia オントロジーは，構築の観点が大きく異なるため，自動的にアライメントをとることは非常に難しいため，本稿では，rdfs:label プロパティを利用しキーワードを登録することにより，手動でアライメントをとった．具体的には，キーワードが登録された動作オントロジーのリソースは，そのキーワードと同名の日本語 Wikipedia オントロジーのリソースに関連付けられる．また，動作オントロジーのクラスにキーワードを登録することにより，それに属する全てのインスタンスに一括してキーワードを登録できる．このようにすることにより，登録コストを減少させることができる．表 1 に，現時点で登録されているキーワードリストの一部を示す．

表 1.　キーワードリスト

動作オントロジーのリソース	キーワード
roboonto_class:移動	移動，動く，歩き，徒歩
roboonto_instance:短い太極拳	太極拳，中国，上海
roboonto_instance:長い太極拳	
roboonto_instance:スリラーダンス	マイケル・ジャクソン，スリラー

これらのキーワードは，protégé*6 などのオントロジーエディタを用いることで容易に編集できる．

8. 対話の流れ，ケーススタディ

WioNA では，日本語 Wikipedia オントロジーを使うことによって，膨大な知識に基づいて対話を進める．対話のスタート地点についても，約 5000 個のトピックからユーザが選択できる．また，対話の各フェーズにおいて，Nao は動作オントロジーを参照し，関連動作がある場合にはそれを実行する．さらに，実行後，動作オントロジーのインスタンストリプルを使って，ユーザからのより短い動作や簡単な動作の要求に応えることも可能である．本稿では，WioNA の有用性を示すために，健康法とロック歌手という 2 つのトピックにおける対話ログを示す．

8.1　ケーススタディ I　Topic：健康法

（ユーザ）健康法について教えて．
（Nao）健康法について調べてみますね．
（Nao）健康法についてはたくさん分かるから，5 個ずつ言うね．健康法には，早寝早起き，太極拳，ラジオ体操，日光浴，禁煙があるよ．この中に何か気になるのあった？それとももっと他のを聞きたい？
（ユーザ）太極拳について教えて．
（Nao）太極拳には，孫式太極拳，陳式太極拳，武式太極拳，呉式太極拳があるよ．この中に何か気になるのあった？
（ユーザ）孫式太極拳について教えて．
（Nao）孫式太極拳に関係ある動作として，短い太極拳と長い太極拳があるよ．どうかな？
（ユーザ）短い太極拳やって．
（Nao）了解しました．短い太極拳をやりますね．

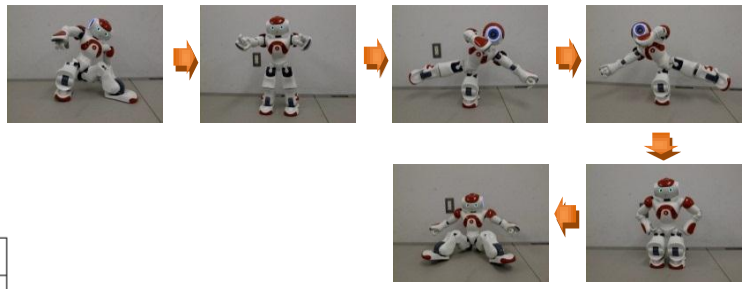

図 6　短い太極拳を実行

（ユーザ）もっと短い動作やって．
（Nao）了解しました．

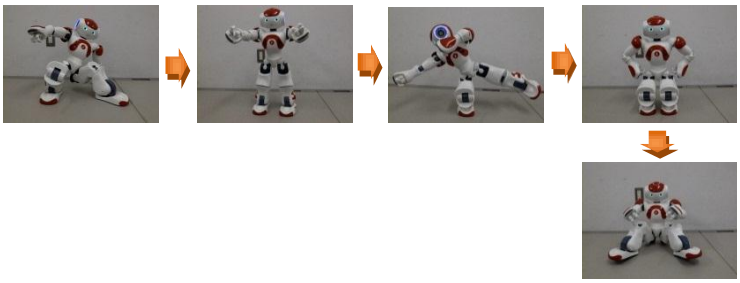

図 7　前提条件に基づき動的に生成した短い太極拳の短縮バージョンを実行

8.2 ケーススタディ II　Topic：ロック歌手

（ユーザ）ロックの殿堂入り受賞者について教えて．
（Nao）分かりました．ロックの殿堂入り受賞者について調べてみますね．
（Nao）ロックの殿堂入り受賞者についてはたくさん分かるから，5 個ずつ言うね．ビリー・ホリデイ，アース・ウィンド・アンド・フ

ァイアー，ロッド・スチュワート，ハンク・ウィリアムズ，ヴェルヴェット・アンダーグラウンド．この中に何か気になるの合った？それとももっと他のを聞きたい？
（ユーザ）マイケル・ジャクソンについて教えて．
（Nao）マイケル・ジャクソンに関連ある動作は，スリラーダンスが出来るからやります．

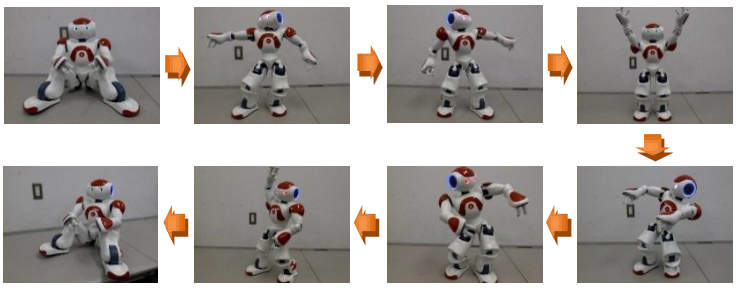

図 8　スリラーダンスを実行

（Nao）他にも何か聞きたい？
（ユーザ）もう終わり．

8.3 ケーススタディにおける考察

2 つのケーススタディは，健康法，ロック歌手という 2 つのトピックにおいて，よりよい HRI を実現できることを表している．しかし，同時にこれらは改善しなければいけない大きな課題を明らかにしている．第一に，現状では，対話はあくまでの日本語 Wikipedia オントロジーのネットワークをたどることによって進められる．そのため，対話の自由度が低いという大きな問題がある．ロボットとの対話の特性上，完全に対話の主導権をユーザ側に渡すということは不可能であるが，ユーザの質問を SPARQL クエリへマッピングすることにより，ユーザから簡単な構文の 5W1H の質問に答えられるようにする必要がある．二つめに，現状では，Nao が実行可能な動作が限られているために，関連動作を実行できるトピックが非常に限られているということである．したがって，非常に幅広い分野における対話が可能だが，その大部分は動作を伴わず，人型ロボットを利用するメリットを生かしきれていないという問題がある．したがって，前提条件を利用してより柔軟に動作を動的に生成するフレームワークを構築する必要がある．

9. 結論と今後の展望

本稿では，日本語音声インタフェース，意味理解，日本語 Wikipedia オントロジー，動作オントロジーの 4 要素を統合することによって，膨大な知識に基づく対話と関連動作の実行を実現するシステム WioNA を提案した．そして，2 つのケーススタディにより，WioNA が健康法，ロック歌手というトピックにおいて，ユーザとの HRI がうまくいくことを示した．今後の展望としては，前述の 2 つの課題に取り組むと共に，ユーザにセンサを装着し，そのフィードバックを利用することにより，対話だけでなく，動作についても双方向のコミュニケーションを実現したいと考えている．また，ユーザのより高度な要求を異機能ロボットの連携により実現する仕組み[7]についても検討していくつもりである．

参考文献

[1] B.Johnston, F.Yang, R.Mendoza, X.Chen, M.Williams: Ontology Based Object Categorization for Robots, Proc. of The 7th International Conference, Practical Aspects of Knowledge Management(PAKM2008), Yokohama Japan (2008)

[2] I.Suh, G.Lim, WHwang, H.Suh, J.Choi, Ontology-based Multi-layered Robot Knowledge Framework (OMRKF) for Robot Intelligence. Proc of 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, (2007)

[3] 川本真一，下平博，新田恒雄，西本卓也，中村哲，伊藤克亘，森島繁生，四倉達夫，甲斐充彦，李晃伸，山下洋一，小林隆夫，徳田恵一，広瀬啓吉，峯松信明，山田篤，伝康晴，宇津呂武仁，嵯峨山茂樹：Open-source Software for Developing Anthropomorphic Spoken Dialog Agent',' Proc. of PRICAI-02, International Workshop on Lifelike Animated Agents, pp.64-69, (2002).

[4] Soren Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann1, Richard Cyganiak, Zachary Ives: DBpedia: A Nucleus for a Web of Open Data, Lecture Notes in Computer Science ,Springer Berlin / Heidelberg ,pp.722-735(2007).

[5] Fabian M.Suchanek,Gjergji Kasneci,Gerhard Weikum: Yago: a core of semantic knowledge, Proceedings of the 16th international conference on World Wide Web, ACM , pp.697-706(2007).

[6] 玉川奨，桜井慎弥，手島拓也，森田武史，和泉憲明，山口高平: Learning a Large Scale of Ontology from Japanese Wikipedia, 2010, IEEE/WIC/ACM International Conference on Web Intelligence, pp279-286

[7] 宮川智好，山本隆三，植田光，山口高平: 多重オントロジーに基づくセマンティックロボットサービスの設計と実現，第 23 回人工知能学会全国大会 (2009).