

オントロジーを用いた英日料理レシピ変換システム

木村美香子[†] 北村 泰彦[†] 松田 匡史[†] YuriTijerino^{††}

[†] 関西学院大学理工学部

^{††} 関西学院大学総合政策学部

E-mail: †{mikako.kimura,ykitamura,bfv59036}@ksc.kwansei.ac.jp, ††yuri@tijerino.net

あらまし Web 上には様々な国のレシピサイトが数多く存在する。海外の料理レシピを用いた料理を自分の国で作る場合、外国語でレシピが書かれているため、料理の作り方が理解できないという問題があった。さらに、たとえレシピ内容を翻訳できたとしても、その国独特の材料や調理法がレシピに記述されているため、自分の国で材料が手に入らないといった問題があり、外国の料理を作ることが困難であった。本研究ではオントロジーを用いて、海外レシピを自分の国で調理可能な内容に変換する手法を提案する。単にレシピを翻訳するのではなく、オントロジーに記述してある材料や調理法の特徴を用いてレシピ変換を行う。キーワード 料理オントロジー, 中間コード, 機械翻訳

English-Japanese Cooking Recipe Translation System Using Ontology

Mikako KIMURA[†], Yasuhiko KITAMURA[†], Masafumi MATSUDA[†], and Yuri TIJERINO^{††}

[†] School of Science and Technology, Kwansei Gakuin University 2-1 Gakuen, Sanda-shi, Hyogo, 669-1337 JAPAN

^{††} School of Policy Studies, Kwansei Gakuin University 2-1 Gakuen, Sanda-shi, Hyogo, 669-1337 JAPAN

E-mail: †{mikako.kimura,ykitamura,bfv59036}@ksc.kwansei.ac.jp, ††yuri@tijerino.net

Abstract It is difficult to cook foods following a recipe on foreign Web sites, because it is difficult to understand the recipe and to get the foreign ingredients listed in it. In the paper, we propose a method to translate foreign recipes into cookable Japanese recipes by using ontology. We do not only translate the recipes into Japanese but also replace foreign ingredients with the corresponding Japanese ones by using ontology.

Key words cooking ontology, intermediate code, machine translation

1. はじめに

Web 上には様々な国のレシピサイトが数多く存在する。例えば”All recipes” [1] などの料理レシピサイトでは、英語で記述された海外の様々な料理レシピが提供されている。海外の料理レシピを用いた料理を自分の国で作る場合、外国語でレシピが書かれているため、料理の作り方が理解できないという問題があった。さらに、たとえレシピ内容を翻訳できたとしても、その国独特の材料や調理法がレシピに記述されているため、自分の国で材料が

手に入らないといった問題がある。例えば、「turkey」を従来の機械翻訳ツールで翻訳すると、「七面鳥」となる。しかし、「七面鳥」は日本では入手が容易でない海外独特の材料である。また「pound」のように、海外独特の単位や「cup」という海外と日本では1cupあたりの量が異なる単位もある。以上のことから、従来の機械翻訳ツールによる料理レシピの翻訳結果だけで外国の料理を作ることが容易ではない。

本研究では料理レシピの特徴を用いて、海外レシピを自分の国で調理可能な内容に変換する手法を提案する。

単にレシピを翻訳するのではなく、材料や調理法の特徴を用いてレシピ変換を行う。材料や調理法の特徴を表現する方法として、オントロジーを用いる。本論文では、以下、2章において機械翻訳による海外レシピ翻訳の問題点と、その解決法について述べる。3章においてレシピページからレシピ情報を抽出し、XML タグでレシピ情報を判別する方法について述べる。次に4章で料理オントロジーと中間コードを用いた、レシピ変換方法について述べる。5章では提案した方法を用いて、実際の海外料理レシピを英文から日本語へと変換した結果について述べる。6章で関連研究について述べ、6章で本論文のまとめとする。

2. レシピ翻訳の問題点

海外の料理レシピを従来の機械翻訳ツールで翻訳する際、以下の2つの問題点が挙げられる。

(1) 料理の専門用語が適切に翻訳されない。

(2) 食材の翻訳結果として、日本で入手が容易でない食材に変換される。

例として、図1のレシピを従来の機械翻訳ツールを用いて翻訳する。この翻訳結果を図2に示す。

```
Ingredients:
22 pounds whole turkey
1/2 teaspoon dried sage
1 cup apricot preserves
.....
Directions:
.....
Season turkey cavity with salt and pepper.
.....
```

図1 レシピ情報の抽出結果例

まず(1)の問題点として、図2では、材料で「ポンド」という単位が記述されているが、日本では使われない単位のため、このままでは正確な量がわからない。また、「1 cup apricot preserves」が「アプリコットが保存する1個のカップ」と翻訳されており、適切な翻訳結果とはいえない。さらに、作り方を見ると、「Season」という単語は料理レシピでは本来「味付けをする」という意味なのに対し「季節」と誤った翻訳がされている。次に(2)の問題点について、翻訳されたレシピ中に記述されている「七面鳥」は日本では入手が容易でない食材であり、日本で手に入る食材を考慮せずに翻訳されている。以上のように、従来の翻訳ツールでは適切に海外料理レシピが翻訳されないため、日本でこの料理レシピを調理することは容易ではない。

本研究では、料理オントロジーを用いることで海外料理レシピ翻訳におけるこれら2つの問題を解決する。まず、

```
材料:
22 ポンドの全体の七面鳥
1/2 ティースプーンの乾いた賢人
アプリコットが保存する1個のカップ
.....
作り方:
.....
塩とコショウがある季節七面鳥空洞。
.....
```

図2 従来の機械翻訳ツールを用いた海外料理レシピの翻訳結果

(1)の問題点を解決するために、中間コードを用いた料理レシピの翻訳を行う。従来の翻訳ツールでは、専門用語の対訳が難しいという問題がある。そこで専門辞書を機械翻訳の辞書とは別に作成し、機械翻訳ツールと連携させることで、専門用語の翻訳を可能にする。レシピ文中の専門用語をいったん中間コードに置換して翻訳し、その中間コードを専門辞書に載っている対訳で翻訳する。専門用語の属性を持った単語を中間コードにするため、まず専門辞書内の専門用語をカテゴリ別に分類する。また、中間コードとしてそれぞれのカテゴリに一致し、かつ機械翻訳ツールが扱える単語を用意する。本研究では料理オントロジーを用いることで、「飲み物」、「肉類」などの食材種類別に定義された中間コードを用いていく。

また、(2)の問題点を解決するために、英日2つの料理オントロジーを用いる。海外料理レシピに海外独特の食材が使われている場合、英語の料理オントロジーからその食材の特徴を調べ、その特徴を基に日本の料理オントロジーから代替材料を探す。

このように料理オントロジーを用いることで、海外料理レシピ翻訳におけるこれら2つの問題を解決する。

3. 英日料理レシピ変換システム

本研究では料理オントロジーを用いた英日料理レシピ変換システムを開発した。図3にシステムの構成図を示す。まず、日本で調理可能な内容へと変換したいレシピページを入力とし、wrapperを用いることでレシピ情報を抽出する。次に、抽出したレシピ情報と食材辞書、単位辞書などを用いて、レシピ情報の解析を行い、それをXML形式で表す。XML形式で表されたレシピ情報の食材から、その食材が日本で入手が容易な材料かどうかを、英日2つの料理オントロジーを用いて調べる。もし入手が容易でない食材の場合、英語の料理オントロジーに定義されているその食材の特徴を用いて、日本料理オントロジーから代替材料を探す。また、海外料理レシピ文中の専門用語を、料理オントロジーに定義されている中間コードで置換し、置換後の文を機械翻訳ツールを用いて翻訳する。さらに、先ほど日本料理オントロジーで探した代替材料と、専門辞書を用いることで、翻訳した文の

表 1 料理レシピのタグ一覧

タグ名	解説
recipe	1つの料理レシピを表す
name	料理名を表す
ingredients	料理レシピの材料1文を表す
quantity	数字の値を表す
measure	量などの単位を表す
food	食材を表す
directions	パート別の作り方を表す
section	作り方の文章1文を表す
verb	動詞を表す

中間コードを元に戻す。その結果、海外料理レシピを日本で調理可能な内容に変換できる。以下、それぞれの詳細を述べる。

3.1 レシピページからの情報抽出

本研究ではまず、海外のレシピサイトにある外国語で書かれた料理レシピページから、レシピ情報を wrapper を用いて抽出する。レシピ情報として、材料と作り方の内容を料理レシピサイトから抽出する。レシピサイトにおける料理レシピ情報は HTML として記述されているため、wrapper を用いることで料理の材料、作り方等の必要な情報を抽出する。wrapper とは、HTML ページにおけるタグの構造などを手掛かりに、Web ページから特定の部分を抽出するためのプログラムである。図 1 にレシピ情報の抽出結果例を示す。

次に、抽出結果から、レシピ中のフレーズを”量”，”単位”，”食材”などに判別していく。例として、図 1 では、食材辞書を用いることで、”turkey”，”sage”，”apricot preserves”が食材のフレーズであることを判別する。同様に、単位辞書を用いることで、”pounds”，”cup”，”teaspoon”が単位のフレーズであることを判別する。また、”1/2”などの分数は、小数”0.5”に変換する。さらに、動詞辞書を用いることで、”season”などが動詞であることを判別する。そして、レシピ情報から「量」，「単位」，「食材」などを判別した結果をそれぞれの種類ごとにタグ付けしていき、XML 形式で表していく。この例では、”12”は量を示す<quantity>，”pounds”は単位を示す<measure>，”turkey”は食材を示す<food>でタグ付けしていく。この結果、量、単位、食材などを判別した結果を、図 4 のような XML 形式で表すことができる。表 1 に XML 形式で料理レシピを表すのに用いるタグについて示す。

3.2 オントロジーを用いたレシピ変換

次に、レシピページから抽出したレシピ情報と、料理オントロジーを用いて、レシピ情報を日本で調理可能な内容に変換していく。

(a) 料理オントロジー

意味情報を表現する方法の一つに、オントロジーがある。オントロジー [2] [3] [4] [5] とは概念同士の相互関係を

```

<recipe>
<name>Ma Lipo's Apricot-Glazed Turkey with
Roasted Onion and Shallot Gravy </name>
<ingredients>
<quantity>22</quantity> <measure>pounds
</measure> whole <food>turkey</food>
</ingredients>
<ingredients>
<quantity>1</quantity> <measure>cup
</measure><food>apricot preserves</food>
</ingredients>
<ingredients>
<quantity>0.5</quantity> <measure>teaspoon
</measure> dried <food>sage</food>
</ingredients>
.....
<directions>
<section>
<verb>Season<verb> <food>turkey</food> cavity
with <food>salt</food> and <food>pepper</food>.
</section>
.....
</directions>
</recipe>

```

図 4 レシピ内容の XML 形式による記述

表したものであり、本研究では料理オントロジー [6] を用いることで、料理レシピの内容を定義する。本研究で扱う料理オントロジーは、料理で用いる食材、調理法などの概念関係を表している。

例として図 5 の料理オントロジーには食材に関する情報が記述されており、左側のオントロジーには肉類”Bird”の種類として”Chicken”，”Turkey”，”Duck”があることを表している。本研究では、外国語の料理レシピを日本語の料理レシピへ変換するために、外国語、日本語の料理オントロジーを用いる。

(b) 料理オントロジーを用いたレシピ変換方法

本研究では、海外料理レシピに書かれた食材が日本で入手が容易でない海外独特のものであった場合、英日 2 つの料理オントロジーを用いて代替材料を探す。英語の料理オントロジーでその食材の特徴を調べ、その特徴を基に日本の料理オントロジーから代替材料を探す。

例として、図 4 の判別結果と、図 5 の外国語、日本語の料理オントロジーを用いて、レシピを変換していく。図 4 では料理レシピの食材として”sage”，”turkey”などがある。まず、これらのキーワードをそれぞれ日本語に翻訳していくと、”sage”の翻訳結果は「セージ」、”turkey”は「七面鳥」である。「セージ」は日本で馴染みのある食

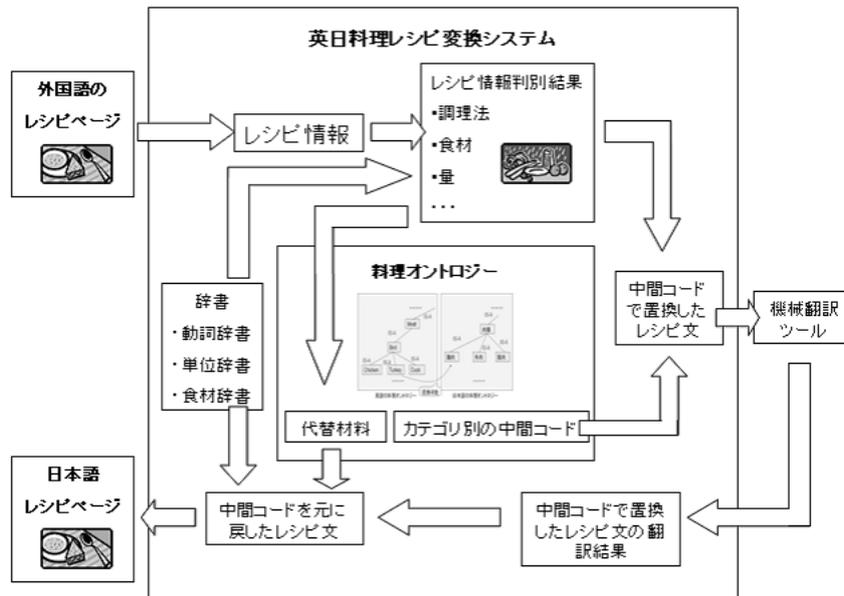


図 3 英日料理レシピ変換システム

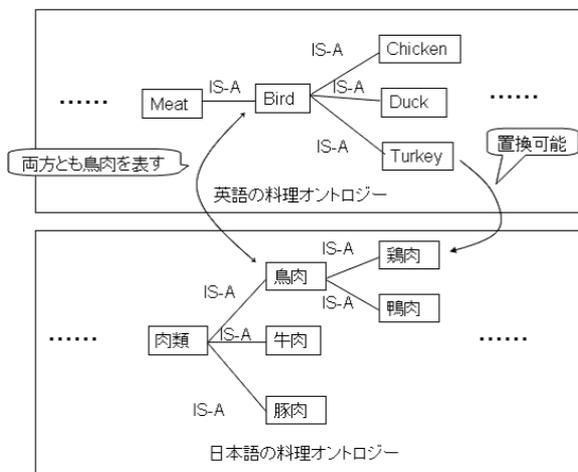


図 5 2つの料理オントロジー

材なので、日本語の料理オントロジーでハーブの一つとして存在する。よって、外国語レシピでの"sage"は、日本語レシピでは「セージ」と変換する。「七面鳥」の場合は、日本ではあまり馴染みのない食材であるため、日本語の料理オントロジーには存在しない。その場合、「七面鳥」の代替材料を日本語の料理オントロジーから探す。外国語オントロジー中において、「Turkey」は鳥肉の一つであると表されているため、「Turkey」の代替材料として日本で手に入る鳥肉を候補とする。日本語の料理オントロジーでは、肉類として「鶏肉」、「牛肉」、「豚肉」があるが、今回は「Turkey」と同じ鳥肉であり、最も一般的に使われる鶏肉を採用する。このように、2つの料理オントロジーを用いることで、「Turkey」の代替材料として鶏肉を推測することが出来る。

(c) 中間コードを用いた英語レシピの翻訳

機械翻訳ツールを用いて海外レシピ文を翻訳する際、従来の機械翻訳ツールでは、専門用語の翻訳が難しいという問題がある[7]。例えば、「Mix sage.」という文を機械翻訳ツールを用いても、「賢人を混ぜてください。」という翻訳結果になり、「sage」という単語がハーブの一種である「セージ」と翻訳されない。そこで、専門辞書を機械翻訳の辞書とは別に作成し、機械翻訳ツールと連携させることで、専門用語の翻訳を可能にする[8]。文中の専門用語をいったん中間コードに置換して翻訳し、その中間コードを専門辞書に載っている対訳で翻訳する。先ほどの例文の場合、中間コードを用いることで、図6のように翻訳される。まず、「Mix sage.」の"sage"を、ハーブの中間コードである"basil"に置き換え、「Mix basil.」とする。この文を機械翻訳ツールで翻訳すると、「バジルを混ぜてください。」に変換される。そして、この文の「バジル」を"sage"の専門辞書に載っている対訳である「セージ」と置換する。その結果、「セージを混ぜてください。」という文となる。

また、料理で用いられる専門的な動詞に関しても、中間コードを用いて翻訳する。動詞に関しても、例えば「Season with salt and pepper.」を翻訳すると、「塩とコショウがある季節」となり、「Season」の適切な翻訳結果である「味付ける」と変換されない。このような場合、食材の場合と同様に、動詞の中間コード「Cook」と「Season」を置換し、「Cook with salt and pepper」という文にする。これを日本語に翻訳すると「塩とコショウで、料理してください。」になる。最後に、中間コードに「Cook」の対訳である「料理して」を、図7の動詞辞書に記述されている「Season」の対訳「味付けて」へと置換する。その結果、

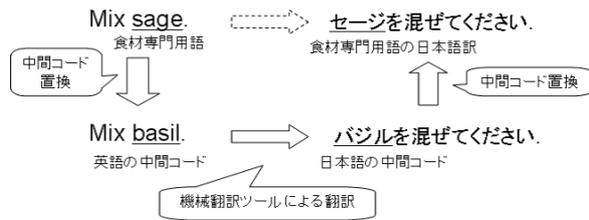


図 6 中間コードを用いたレシピ文の翻訳過程

simmer, 煮る, 煮て
 blend, 混ぜる, 混ぜて
 saute, 炒める, 炒めて
 stir, かき混ぜる, かき混ぜて
 slice, 薄く切る, 薄く切つて
 cream, クリームを加えて調理する, クリームを加えて調理して
 puree, ピューレにする, ピューレにして
 combine, 混ぜる, 混ぜて
 sprinkle, ふりかける, ふりかけて
 beat, 泡立てる, 泡立てて
 microwave, レンジ加熱する, レンジ加熱して
 drain, 水分抜きする, 水分抜きて
 grease, 油をひく, 油をひいて
 season, 味付ける, 味付けて

図 7 専門動詞辞書の例

「塩とコショウで、味付けてください。」に変換される。

専門用語の属性を持った単語を中間コードにするため、まず専門辞書内の専門用語をカテゴリ別に分類する。中間コードとして、それぞれのカテゴリに一致し、かつ機械翻訳ツールが扱える単語を用意する。本研究では料理オントロジーを用いて、「飲み物」、「肉類」、「ハーブ」などの食材種類別に中間コードを定義する。例として、食材「Herb」の中間コードを料理オントロジーに定義すると、図 8 のようになる。「sage」の中間コードを調べる場合、図 8 を見ると「sage」の上位概念として「Herb」が定義されている。「sage」は「Herb」の一種なので、図 8 から、中間コードは「Herb」の代表的なものである「basil」、「thyme」、「dill」の 3 つになる。中間コードが複数存在する理由は、文中に同じカテゴリの専門用語が複数出てきた場合、中間コードが重ならないようにするためである。このように料理オントロジーを用いることで、食材の種類別に分類された中間コードを用いることができる。

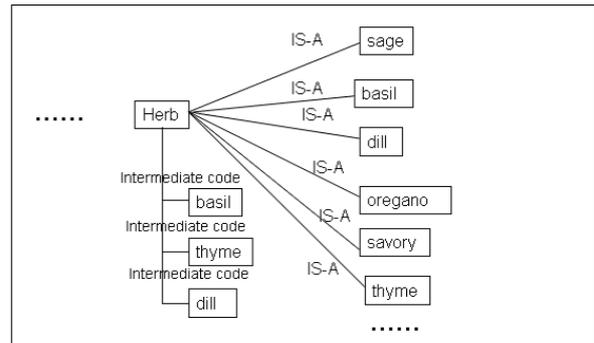


図 8 オントロジーによる中間コードの定義

cup, カップ, 1:1.2
 gallon, リットル, 1:3.8
 ounce, グラム, 1:30
 pound, グラム, 1:450
 quart, ミリリットル, 1:950
 tablespoon, 大さじ, 1:1
 teaspoon, 小さじ, 1:1

図 9 単位辞書の例

(d) 単位と値の変換

海外料理レシピには、「ounce」、「pound」といった日本では使われない単位が使われている。また、日本でも使われている「cup」という単位に関しても、海外での 1cup は日本での 1.2cup と同値であり、海外と日本では 1cup あたりの量が異なる。従来の機械翻訳ツールはそのことを考慮せずに料理レシピを翻訳するため、翻訳結果だけでは海外料理を作ることは困難であった。そこで、本研究では図 9 に示す単位辞書を用いて、海外と日本の単位を考慮した料理レシピの翻訳を行う。単位辞書には英語の単位、日本で用いる代わりにの単位、そしてこれら 2 つの単位を変換する際の値の比率が記述されており、これらを用いて海外と日本における単位の違いを考慮した変換ができる。

例えば、レシピ中に「22 pounds」というフレーズがあったとする。この場合、単位辞書の「pound」の行を見ると、「グラム」というフレーズが書かれているため、日本では「pound」の代わりに「グラム」という単位を用いる。また、「1:450」という比率は、「1 pound」は「450 グラム」と同値であることを表している。よって辞書を用いることで、 $22 \times 450 = 9900$ と計算し、「22 pounds」は「9900 グラム」と変換できる。

4. 関連研究

料理レシピの内容をオントロジーを用いて処理する研

究は数多くある。安枝 [6] の研究では、個人の料理レシピを容易に管理し、食材の情報検索できる Cooking Recipe Searching System の研究を行っている。この研究では、料理オントロジーを用いて、食材、味覚などの抽象的概念の検索することで、自分の管理するレシピを自分好みの内容にすることを目的としている。また、Leticia Gutierrez Villarias [9] は、Web ページ上にある料理レシピからレシピ情報を抽出するアノテーションツールの研究を行っている。この研究では、料理オントロジーに定義された食材、単位の特徴を用いることで、料理レシピからレシピ情報を抽出することを目的としている。本研究においても料理オントロジーは用いているが、海外の料理レシピを自分の国で調理可能な内容に変換するという目的という点で、前述した 2 つの研究と異なる。

増田ら [8] は、中間コードを用いて機械翻訳を行う、専門辞書拡張型機械翻訳システムの研究を行っている。本研究でも中間コードを用いた機械翻訳を行っているが、オントロジーでカテゴリ別の中間コードを定義している点でこの研究と異なる。

5. まとめと今後の課題

本研究では料理オントロジーを用いることで、海外料理レシピを日本でも調理可能な内容への変換することを目的とした。料理オントロジーに記述された食材の特徴を用いることで、海外独特の食材を日本でも手に入る代替材料へと変換した。また、料理オントロジーに定義した食材種類別の中間コードを用いることで、従来の翻訳ツールでは困難だった料理専門用語の翻訳を可能にした。

今後の課題として、英日料理レシピ変換システムと Semantic Wiki [10] との連携があげられる。英日料理レシピ変換システムによる海外レシピの変換結果について、より適した材料をユーザにレシピに入力させることで、よりユーザ好みの内容に料理レシピを編集することができる。例えば英日料理レシピ変換システムが、海外料理レシピに書かれた「七面鳥」の代わりに「鶏肉」を推薦したとする。しかし、ユーザは「鶏肉」よりも「鴨肉」の方がこのレシピにより適していると判断したとする。そこでユーザが Semantic Wiki を用いることで、海外レシピの変換結果を「鶏肉」から「鴨肉」へと編集し、システムにユーザの編集内容を学習させる。これにより、今までのレシピ変換ではシステムは「七面鳥」を「鶏肉」と変換していたが、次のレシピ変換から「鶏肉」ではなく「鴨肉」へと変換することができ、ユーザの編集内容を反映したレシピ変換を行うことができる。

また、もう一つの課題として、レシピに書かれた食材を、ユーザが Semantic Wiki を用いて編集することで、食材の編集内容をレシピの作り方に反映させることが挙げられる。例えば、システムによる海外レシピの変換結果に対し、ユーザが「鶏肉」から「鴨肉」へと食材を修

正したとする。するとシステムは料理オントロジーを用いて、レシピの作り方に書かれた煮込み時間、下準備などを鶏肉用から鴨肉用のものへと変換する。このように、ユーザによる食材の編集内容に応じた調理法へとシステムが変換してくれるため、ユーザの編集内容に応じてレシピを発展させることができる。

文 献

- [1] All recipes, <http://allrecipes.com/>
- [2] 神崎 正英. セマンティック・ウェブのための RDF/OWL 入門, 森北出版株式会社, 2005.
- [3] 情報処理相互運用技術協会編:セマンティック Web 入門. オーム社, 2004.
- [4] Mikako Kimura and Yasuhiko Kitamura. Embodied Conversational Agent Based on Semantic Web. Zhong-Zhi Shi and Ramakoti Sadananda (Eds.): Agent Computing and Multi-Agent Systems, Lecture Notes in Artificial Intelligence 4088, Springer-Verlag, pp.734-741, 2006.
- [5] 溝口理一郎. オントロジー構築入門, オーム社, 2006.
- [6] 安枝明日香, 料理分野におけるオントロジー. 関西学院大学総合政策学部卒業論文, 2007.
- [7] Toru Ishida. Language Grid: An Infrastructure for Intercultural Collaboration. In IEEE/IPSJ. Symposium on Applications and the Internet(SAINT-06), pp.96-100, 2006.
- [8] 増田雄介, 北村泰彦. 専門辞書拡張型機械翻訳システムにおける中間コード選択, 第 51 回システム制御情報学会研究発表講演会, 4F3-2, 2007.
- [9] Leticia Gutierrez Villarias. Ontology-Based Semantic Querying of the Web with Respect to Food Recipes. Master's thesis, Department of Informatics and Mathematical Modelling, Technical University of Denmark, 2004
- [10] Kensaku Kawamoto, Yasuhiko Kitamura and Yuri Tjjerino. KawaWiki: A Template-Based SemanticWiki Where End and Expert Users Collaborate, 5th International Semantic Web Conference, Poster paper, 2006.