

## 料理提案システムのための料理概念ベースの構築

森泰宏<sup>†1</sup> 吉村枝里子<sup>†2</sup> 土屋誠司<sup>†2</sup> 渡部広一<sup>†2</sup>

コンピュータは人間のようにある料理から関連のある他の料理を連想することは不可能である。料理の関連性を理解させるためには、料理の類似度を定量的に評価する必要がある。そこで本稿では料理概念ベースを構築する。料理概念ベースはある語（概念）に対し、その特徴を表す語（属性）と属性の重要性（重み）の対から構成されている。料理概念ベースを構築することで料理の類似度を定量的に評価することを目指す。

### Construction of the Cooking Concept Base for Cooking Recommendation System

YASUHIRO MORI<sup>†1</sup> ERIKO YOSHIMURA<sup>†2</sup>  
SEIJI TSUTHIYA<sup>†2</sup> HIROKAZU WATABE<sup>†2</sup>

Computer can't associate other cooking with one cooking like humans. It is necessary to evaluate the similarities of cooking in order to understand the relevance between cooking. So we construct the Cooking Concept Base in the present study. The Cooking Concept Base is constructed from a word (concept) and words (attribute) representing the features of the concept and values (weight) that represents the importance of the attribute. We evaluate the similarities between cooking by constructing the Cooking Concept Base.

#### 1. はじめに

料理は人が生きていく上で非常に重要な役割を果たしている。また、人の健康を支える源としても大きな役割を果たしている。しかし、毎日料理を作る人（主婦など）にとって、献立を決めることは大きな負担になる場合がある。料理を作る人が献立に困ったときは、料理本やレシピサイトを用いて献立を決めることが多い。しかし、レシピサイトで料理を検索すると、すべてのユーザに同じ結果を返すだけで、ユーザはその膨大な量のレシピから作りたい料理を探さなければならない。

そこで料理を提案するシステム（料理提案システムとする）の必要性が考えられる。料理提案システムはユーザの入力（好きな食べ物・材料や嫌いな食べ物・材料）から料理を提案するシステムである。一つの手法として、ユーザの入力（材料）と同じ材料を使用している料理を提案することが挙げられる。しかし、この手法では、料理を使用材料の有無のみで判断するため、料理を提案するにあたって、ユーザの知らない料理や食材を提案する可能性は低い。しかし、人間であれば、「レモン」が好きであれば、「ゆず」も好きかもしれない、更に「ゆず茶」を好むかもしれない、などを連想することにより新たな見地から料理を提案することができる。

そこでこのような機構をシステムに組み込むため、料理

にある材料があるかないかで判断するのではなく、ある材料がどれだけその料理に関連しているかを表す関連度を考慮する必要がある。

自然言語において、語と語の関連性を定量的に評価する手段として関連度計算方式[1]が提案されている。関連度計算方式は語の特徴を辞書や新聞を用いて格納した概念ベース[2]を用いて計算を行う。概念ベースと関連度計算方式については2章で詳しく説明する。

提案されている概念ベースは語の一般的な特徴をとらえようとする試みであるが、専門的な用語や、その分野における特別な関連性を表現するには弱い。そこで、料理という分野における特別な関連性を表現するために、この概念ベースの考え方を料理という分野に応用した料理概念ベースを本稿で構築する。料理概念ベースには料理に関する料理名や材料などの語（概念）ごとにそれを特徴づける語を属性として付与し、各々の属性には関連の深さを定量的に表現する重みを付与する。属性は概念の意味特徴を表現する役割を有し、属性の集合体によって概念一つが持つ知識を構成している。料理概念ベースを構築することで、料理の類似度を計算し、定量的に評価することが可能になる。料理の類似度とは料理同士の共通する材料の類似度を示す。

本稿ではコンピュータが関連のある料理の連想を可能にするために料理概念ベースの構築を行うことを目的とする。料理概念ベースを構築することで、料理の類似度を定量的に評価することが可能になる。

<sup>†1</sup> 同志社大学大学院 理工学研究科  
Graduate School of Science and Engineering, Doshisha University

<sup>†2</sup> 同志社大学 理工学部  
Faculty of Science and Engineering, Doshisha University

## 2. 概念ベースと関連度計算方式

### 2.1 概念ベース

概念ベースとは、電子化された国語辞書や新聞記事などからある一定のルールに従って自動的に構築された知識ベースである。見出し語（概念）に対して、その特徴を表す語（属性）および属性の重要さ（重み）の対を複数付与することで構成されている。ある概念  $A$  は  $m$  個の属性  $a_i$  と重み  $w_i (>0)$  の対によって次のように表現される。

$$\text{概念 } A = \{(a_1, w_1), (a_2, w_2), \dots, (a_m, w_m)\} \quad (1)$$

概念  $A$  の意味定義を行う属性  $a_i$  を、概念  $A$  の 1 次属性と呼ぶ。概念ベースの特徴として、属性を成す単語群も概念ベースの中で概念として定義されている点がある。つまり、属性  $a_i$  を概念とみなして更に属性を導くことができる。概念  $a_i$  から導かれた属性  $a_{ij}$  を、元の概念  $A$  の 2 次属性と呼ぶ。概念ベースの具体的な例を表 1 に、概念ベースの構造を図 1 に示す。

表 1 概念ベースの例

概念	属性
医者	(医師,0.34)(患者,0.11)(病院,0.08)・・・
病院	(医院,0.25)(手術,0.18)(施設,0.04)・・・
治す	(治療,0.43)(医療,0.21)(病気,0.13)・・・
⋮	⋮

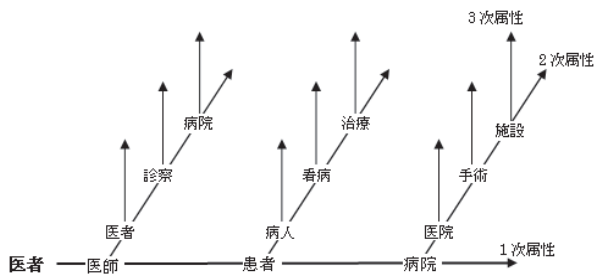


図 1 概念ベースの構造

図 1 に示した通り、ある概念からは任意の次元までの属性を 1 次、2 次、3 次、 $\dots$ 、 $N$  次と導くことができる。このことより概念ベースは、 $N$  次の属性による連鎖構造によって作られていると言える。この連鎖構造によって一つの概念を大量の語で表現することが可能となる。また概念と概念の関連度計算方式を行うことが可能になる。

#### 2.1.1 概念ベースの構築方法

国語辞書からの概念ベースの構築方法を以下に述べる。見出し語の説明文は見出し語の意味を説明している。そのため見出し語を概念としたとき、その説明文にある各語は見出し語の意味特徴を表す語、つまり属性と言える。よって、見出し語からは概念を獲得し、見出し語の説明文から

は属性を獲得することで概念ベースを構築する。

概念ベースに定義される概念は、コンピュータが常識的な判断を行うために概念を連想する際必要な語であるため、それ自身が単独で意味を持つ語である。語自身が単独で意味を持つ語を自立語と呼び、「名詞」、「形容詞」、「動詞」を以下自立語と定義する。

次に、見出し語の説明文を形態素解析することにより、属性になる自立語を抽出する。見出し語の説明文に対して茶筌[3]による形態素解析を行い、その結果が「名詞」、「形容詞」、「動詞」となるものを属性として選択する。

#### 2.1.2 属性の重み付け

属性に重みを付与する手法の 1 つとして、概念ベース  $tf \cdot idf$  法を用いる。概念ベース  $tf \cdot idf$  法とは、文書処理におけるキーワードの重み付け手法として広く利用されている  $tf \cdot idf$ [4] の考え方を概念ベースに適用したものである。

$tf$  とは、概念ベース内における各概念の頻度を表す値である。概念ベースを仮想的な文書集合として捉えることで算出する。具体的には、概念ベースでは、各概念を  $n$  次元の属性連鎖集合によって定義している。したがって、 $n$  次属性空間内において対象となる概念  $A$  が出現する頻度  $tf_n(A)$  を算出する。例えば、概念「走る」の 2 次属性空間内頻度を求める場合について述べる。概念「自動車」が図 2 に示すような属性を持つ場合、概念「自転車」における 2 次属性空間内頻度は 3 という値になる。同様にすべての概念に対して「走る」の 2 次属性内頻度を求め、その合計値を  $tf_2(\text{走る})$  の値とする。

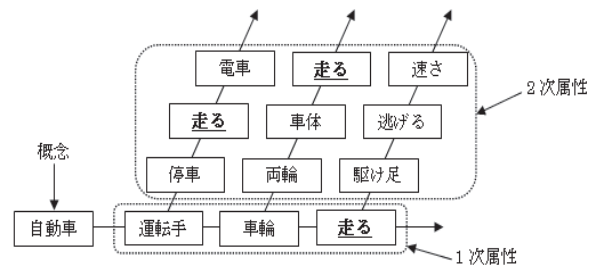


図 2 概念「自動車」における属性「走る」の頻度

$idf$  とは、概念ベース内における各概念の特異性を表す値である。 $n$  次属性空間内において、対象となる概念を属性として持つ概念の総数を算出することができる。頻出する語に比べて、稀に出現する語は概念をより特徴づけることができると考えられる。例えば、概念「人」を概念ベースの  $n$  次属性空間内で属性として持つ概念は多いが、概念「人力車」を  $n$  次属性空間内で持つ概念は少ない場合、この文書空間内では概念「人力車」の方が概念「人」より概念を特徴付けることができる。 $n$  次属性空間内における概念  $A$  の概念価値として、 $idf_n(A)$  を(2)式によって定義する。 $V_{all}$  は概念ベースに定義されている概念数、 $df_n(A)$  は  $n$  次属性空間内で概念  $A$  を属性として持つ概念の数である。以上の  $tf$

と  $idf$  の値を利用し、ある概念  $A$  の属性  $a_i$  の重み  $w(A, a_i)$  を以下の式によって与える。

$$idf_n(A) = \log_2 \frac{V_{all}}{df_n(A)} \quad (2)$$

$$w(A, a_i) = tf_n(a_i) \times idf_n(A) \quad (3)$$

## 2.2 関連度計算方式

関連度計算方式とは、概念ベースに定義されている二つの概念間の関連の強さを定量的に表現する手法である。関連度は 0.0 から 1.0 の間の実数値で表され、概念間の関連が強いほど大きな値を示す。例えば、概念「自動車」に対して、「車」、「自転車」、「馬」の関連の強さを、表 2 に示すように数値化でき、コンピュータにも「自転車」と関連がより強いのは「車」であると判断できるようになる。

表 2 関連度計算方式の例

基準概念	対象概念	関連度
自動車	車	0.4
	自転車	0.18
	馬	0.02

### 2.2.1 一致度

ある概念  $A, B$  において、その属性を  $a_i, b_j$ 、対応する重みを  $u_i, v_j$  とし、それぞれ属性が  $L$  個、 $M$  個 ( $L \leq M$ ) とすると、概念  $A, B$  はそれぞれ

$$A = \{(a_1, u_1), (a_2, u_2), \dots, (a_L, u_L)\} \quad (4)$$

$$B = \{(b_1, v_1), (b_2, v_2), \dots, (b_M, v_M)\} \quad (5)$$

となる。このとき、概念  $A$  と概念  $B$  の属性一致度  $DoM(A, B)$  を以下のように定義する。

$$DoM(A, B) = \sum_{a_i=b_j} \min(u_i, v_j) \quad (6)$$

ここで、 $a_i=b_j$  は属性同士が一致した場合を示している。つまり、一致度とは概念  $A$  と概念  $B$  それぞれの属性の中で一致したもの内、重みの小さい方の総和となる。これは、小さい方の重みは互いの属性の重みの共通部分となっているので、概念  $A$  と概念  $B$  どちらにも有効な重みだと言えるためである。

### 2.2.2 関連度の算出

関連度は次の手順で求める。2.2.1 節で述べた概念  $A, B$  において、まず属性数の少ない方の概念  $A$  を基準とし、その属性の並びを固定する。その上で概念  $B$  の属性を概念  $A$  の各属性との一致度の和が最大になるように並び替える。このときの概念  $B$  の属性と重みを  $(b_{xi}, v_{xi})$  として次のように定義する。

$$B = \{(b_{x1}, v_{x1}), (b_{x2}, v_{x2}), \dots, (b_{xM}, v_{xM})\} \quad (7)$$

これらの概念についての関連度  $DoA(A, B)$  を次の式で定義する。

$$DoA(A, B) = \sum_i DoM(a_i, b_{xi}) \times \frac{(w_i + v_{xi})}{2} \times \frac{\min(w_i, v_{xi})}{\max(w_i, v_{xi})} \quad (8)$$

## 3. 料理概念ベース

本章では本稿で構築する料理概念ベースの構築手法について述べる。料理概念ベースは料理、材料（概念）に対して、共起する料理、材料（属性）および属性の重要さ（重み）の対を複数付与することで構成されている。料理概念ベースの具体的な例を表 3 に示す。

表 3 料理概念ベースの具体例

概念	属性、重み
カレー	(カレー,1.0)(鶏肉,0.37)(玉ネギ,0.35)...
シチュー	(シチュー,1.0)(ジャガイモ,0.52) (人参,0.45)...
オムライス	(オムライス,1.0)(玉ネギ,0.37) (マッシュルーム,0.35)...
⋮	⋮

### 3.1 情報源

料理概念ベースを構築するための情報源として、NTT ソーラス[5]のリーフから料理である語を目視によって判断し、カテゴリとして格納する。また格納した料理の材料を E・レシピ[6]から収集を行う。E・レシピとは京都アム有有限会社が提供している料理レシピサイトであり、現在約 2 万件のレシピが登録されている。E・レシピでは一つの料理に対し、料理名、材料、材料の分量が記載されている。それらのデータの中から材料を機械的に集めて登録する。取得したカテゴリと材料の一部を以下に示す(表 4)。

表 4 取得したカテゴリと材料の例

カテゴリ	材料
カレー	鶏肉,玉ネギ,人参,...
シチュー	牛肉,じゃがいも,玉ネギ,...
⋮	⋮

### 3.2 料理概念ベースの構築手法

概念の格納手法と属性の格納手法について述べる。まず概念の格納手法として、3.1 節で取得したカテゴリを抽出し、それを概念として格納する。次に属性の格納方法として、カテゴリと対応している材料を属性として格納する。また料理概念ベースを構築する際、属性は概念として存在するという構造をとるため、概念の追加を行う必要がある。

そこで先ほど属性として格納した材料を概念に追加する。次に追加した概念の属性を格納する。属性の追加方法として、対象の概念と同じ表記の語が登録されているカテゴリと材料の中で共起している語を属性として格納する。

### 3.3 属性の重み付け

料理概念ベースの属性に、その重要性を意味する重みを付与する。属性に重みを付与する手法として、料理概念ベース  $tf \cdot idf$  法を用いる。料理概念ベース  $tf \cdot idf$  法とは、2.1.2 節で説明した概念ベース  $tf \cdot idf$  法を料理概念ベースに適用したものである。

$tf$  とは、料理概念ベース内における各概念の頻度を表す値である。料理概念ベースを仮想的な文書集合として捉えることで算出する。例えば、概念「人参」の1次属性空間内頻度を求める場合について述べる。概念「カレー」が図3に示すような属性を持つ場合、概念「カレー」における1次属性空間内頻度は3という値になる。同様にすべての概念に対して「人参」の1次属性内頻度を求め、その合計値を  $tf_i(\text{人参})$  の値とする。

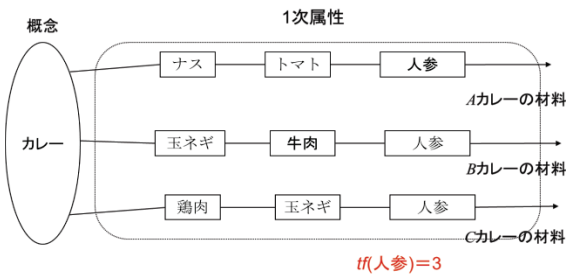


図3 概念「カレー」における属性「人参」の頻度

$idf$  とは、料理概念ベース内における各概念の特異性を表す値である。1次属性空間内において、対象となる概念を属性として持つ概念の総数を算出することができる。頻出する語に比べて、稀に出現する語は概念をより特徴づけることができると考えられる。例えば、概念「塩」を料理概念ベースの1次属性空間内で属性として持つ概念は多いが、概念「人参」を1次属性空間内で持つ概念は少ない場合、この文書空間内では概念「人参」の方が概念「塩」より概念を特徴付けることができる。1次属性空間内における概念  $A$  の概念価値として、 $idf_i(A)$  を(9)式によって定義する。 $V_{all}$  は料理概念ベースに定義されている概念数(3263 語)、 $df_i(A)$  は  $n$  次属性空間内で概念  $A$  を属性として持つ概念の数である。以上の  $tf$  と  $idf$  の値を利用し、ある概念  $A$  の属性  $a_i$  の重み  $w(A, a_i)$  を以下の式によって与える。

$$idf_i(A) = \log_2 \frac{V_{all}}{df_i(A)} \quad (9)$$

$$w(A, a_i) = tf_i(a_i) \times idf_i(a_i) \quad (10)$$

## 4. 料理関連度計算方式

料理関連度計算方式とは、料理概念ベースに定義されている二つの概念間の関連の強さを定量的に表現する手法である。関連度は 0.0 から 1.0 の間の実数値で表され、概念間の関連が強いほど大きな値を示す。例えば、概念「カレー」に対して、「シチュー」、「パスタ」、「ラーメン」の関連の強さを、表5に示すように数値化し、コンピュータにも「カレー」と関連がより強いのは「シチュー」であると判断できるようになる。

表5 関連度計算方式の例

基準概念	対象概念	関連度
カレー	シチュー	0.62
	パスタ	0.25
	ラーメン	0.20

### 4.1 一致度

ある概念  $A, B$  において、その属性を  $a_i, b_j$ 、対応する重みを  $u_i, v_j$  とし、それぞれ属性が  $L$  個、 $M$  個 ( $L \leq M$ ) とすると、概念  $A, B$  はそれぞれ

$$A = \{(a_1, u_1), (a_2, u_2), \dots, (a_L, u_L)\} \quad (11)$$

$$B = \{(b_1, v_1), (b_2, v_2), \dots, (b_M, v_M)\} \quad (12)$$

となる。このとき、概念  $A$  と概念  $B$  の属性一致度  $DoM(A, B)$  を以下のように定義する。

$$DoM(A, B) = \sum_{a_i=b_j} \min(u_i, v_j) \quad (13)$$

ここで、 $a_i=b_j$  は属性同士が一致した場合を示している。つまり、一致度とは概念  $A$  と概念  $B$  それぞれの属性の中で一致したもの内、重みの小さい方の総和となる。これは、小さい方の重みは互いの属性の重みの共通部分となっているので、概念  $A$  と概念  $B$  どちらにも有効な重みだと言えるためである。

### 4.2 料理関連度の算出

料理関連度は次の手順で求める。4.1 節で述べた概念  $A, B$  において、まず属性数の少ない方の概念  $A$  を基準とし、その属性の並びを固定する。その上で概念  $B$  の属性を概念  $A$  の各属性との一致度の和が最大になるように並び替える。このときの概念  $B$  の属性と重みを  $(b_{X1}, v_{X1})$  とし、次のように定義する。

$$B = \{(b_{X1}, v_{X1}), (b_{X2}, v_{X2}), \dots, (b_{XM}, v_{XM})\} \quad (14)$$

これらの概念についての料理関連度  $DoA(A, B)$  を次の式で定義する。

$$DoA(A, B) = \sum_i DoM(a_i, b_{xi}) \times \frac{(w_i + v_{xi})}{2} \times \frac{\min(w_i, v_{xi})}{\max(w_i, v_{xi})} \quad (15)$$

## 5. 評価

料理概念ベースの有用性を示すために、料理概念ベースを用いた料理提案システムと用いないシステムとの比較実験を行う。

### 5.1 料理概念ベースを用いた料理提案システム

料理概念ベースを用いた料理提案システムは好きな食べ物、嫌いな食べ物の嗜好情報から連想を行い、料理を提案するシステムである。

システムの流れを以下に説明する(図4)。まず、ユーザは嗜好情報として好きな食べ物、嫌いな食べ物を任意に入力する。次に、好きな食べ物、嫌いな食べ物のそれぞれから条件に合う料理を料理知識ベースから取得する。次に取得した料理の順位付けを行い、提案する料理を決定する。

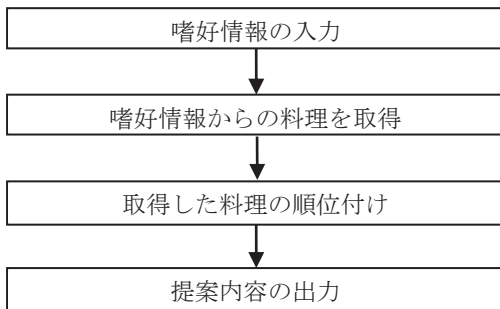


図4 システムの流れ

料理知識ベースとはE・レシピから取得した料理を1596語格納した知識ベースである。料理名それぞれにNTTセンサーラスから取得したカテゴリ(163語)と料理に使用されている材料を付与する。なお、材料の全種類は1506種類となった。料理知識ベースの一部を以下に示す(表6)。

表6 料理知識ベース

料理名	カテゴリ	材料
チキンカレー	カレー	鶏肉,玉ネギ,人参,...
ベジタブルカレー	カレー	ナス,レンコン,人参,...
ビーフシチュー	シチュー	牛肉,ジャガイモ,玉ネギ,...
⋮	⋮	⋮

#### 5.1.1 嗜好情報から料理を取得

嗜好情報から料理を取得する処理の説明を述べる。入力する嗜好情報(好きな食べ物、嫌いな食べ物)には、単なる料理名だけでなく、「辛いもの」といった感覚や「卵」などの材料が入ると考えられる。そこで嗜好情報を「感覚」「料理名・材料」の2パターンに分類する。

まず嗜好情報に形態素解析を行う。形態素解析を行うツールとして茶筌を用いる。例として嗜好情報「甘いもの」に形態素解析を行った結果を表7に示す。

表7 形態素解析結果

単語	読み仮名	活用前の単語	品詞
甘い	アマイ	甘い	形容詞-自立
もの	モノ	もの	名詞-一般

形態素解析の結果、品詞が形容詞と出力された単語が1つでも存在した場合、以下の手法で調味料名を取得する。これ以外の場合は「料理名・材料」とする。

形容詞が存在した場合、調味料知識ベースから一致する形容詞(感覚語)を持つ調味料を全て取得する。調味料を取得する例を図5に示す。調味料知識ベースとは料理知識ベースの材料にある調味料を格納した知識ベースで、それぞれに目視で判断した感覚語を付与している(表8)。

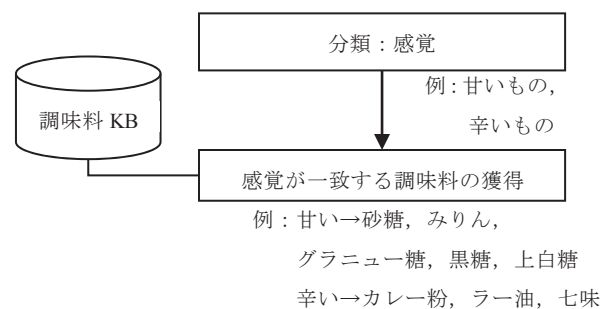


図5 調味料の取得

表8 調味料知識ベース

調味料	感覚
砂糖	甘い
みりん	甘い
レモン	すっぱい
酢	すっぱい
ラー油	辛い

#### 5.1.2 取得した料理の順位付け

取得した料理の順位付けを行うために、4章で述べた料理関連度計算方式を用いる。まず5.1.1節で好きな食べ物から取得した「料理名・材料・調味料」と料理知識ベースのカテゴリとの料理関連度計算方式を行い、その平均値を算出する。これを関連度Aとする。次に嫌いな食べ物から取得した「料理名・材料・調味料」と料理知識ベースのカテゴリとの料理関連度計算方式を行い、その平均値を算出する。これを関連度Bとする。そして関連度Aの値から関連度Bの値を引いた値をその料理名の重みとする。重みの値が正の数のものを提案候補とし、重みが高い順で提案する。以下の図6に例を示す。

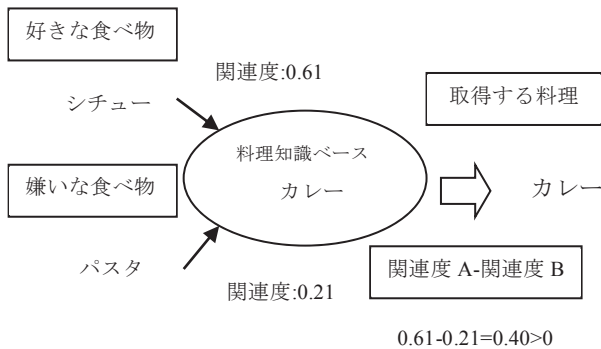


図6 嗜好情報から料理の提案

### 5.2 料理概念ベースを用いない料理提案システム

料理概念ベースを用いない料理提案システムでは、好きな食べ物から「料理」、「材料」、「調味料」に分類する処理は料理概念ベースを用いた料理提案システムと同じである。好きな食べ物が「料理」に分類された場合、料理知識ベースのカテゴリの項目から一致する料理名を取得する。「材料」、「調味料」に分類された場合、料理知識ベースから同じ材料、調味料を含む料理名を取得する。取得した料理を提案候補とする。

次に、嫌いな食べ物から同様に、分類し、それぞれに対し料理知識ベースを参照して料理を取得する。取得した料理が提案候補に存在する場合は、その料理を省き、それ以外の料理を提案する。

### 5.3 評価方法

被験者 10 名にシステムを使用してもらい、実験を行った。また、被験者の 10 名は、嗜好情報として、好きな食べ物と嫌いな食べ物を 3 つまで入力し、出力結果を下記に記す。○、△、×で評価を行った。

- 食べたいものを提案している
- △ どちらかと言えば食べたいものを提案している
- × 食べたくないものを提案している

好きな食べ物と嫌いな食べ物を 3 つ以内で入力する理由は、4 つ以上であると入力の手間がかかるため処理時間が長くなり、被験者の負担が増えると考えたためである。

### 5.4 評価結果と考察

料理概念ベースを用いた料理提案システムでは被験者 10 名に提案した料理の中で 5.3 節の評価方法で○、△、×がついた個数を（被験者の人数×提案した料理の個数）で割った値に 100 をかけた値を精度とした。また重み上位 1 件、2 件、3 件まで提案する 3 パターンの評価をとった。例えば、上位 3 件まで提案した場合、○、△、×の個数を（10×3）で割った値に 100 をかけた値を精度とする。評価の結果、重み上位 3 件の提案した料理の精度は○が 70%、△が 16.7%、×が 13.3%となった。上位 2 件では○が 75%、△が 15%、×が 10%となった。また上位 1 件では○が 80%、△が 10%、×が 10%となった（図 7）。

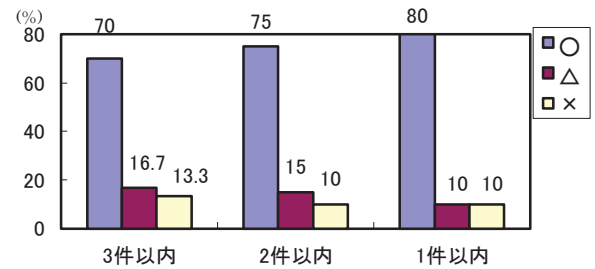


図7 料理概念ベースを用いた提案システムの精度評価

次に、料理概念ベースを用いない料理提案システムでは 6.3 節の評価方法で○、△、×がついた各々の個数を提案した料理の総数で割った値に 100 をかけた値を精度とする。料理概念ベースを用いない料理提案システムで精度評価を行ったところ、○が 59.5%、△が 23.8%、×が 16.7%となった（図 8）。

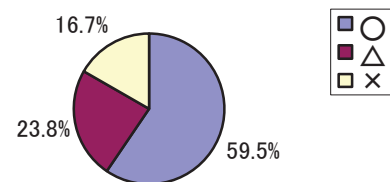


図8 料理概念ベースを用いない提案システムの精度評価

最も多く出現した結果の例として、好きな食べ物が「料理」の場合そのまま同じ料理名を提案する結果が挙げられる。この理由として、料理概念ベースを用いない料理提案システムの場合、好きな食べ物「料理」と料理知識ベースの表記で取得しているためである。一方、料理概念ベースを用いた料理提案システムでは同じ料理同士で料理関連度計算方式を行うとき、関連度の値が必ず 1 になる。そのため重み付けで関連度の差をとっても値が大きくなってしまった。これは単に同じ料理を提案しているだけであり、本稿の特徴である、ある料理から違う料理を連想することができていないので、入力と同じ料理を提案した場合、その料理を評価に入れないパターンの精度をとった（図 9、図 10）。

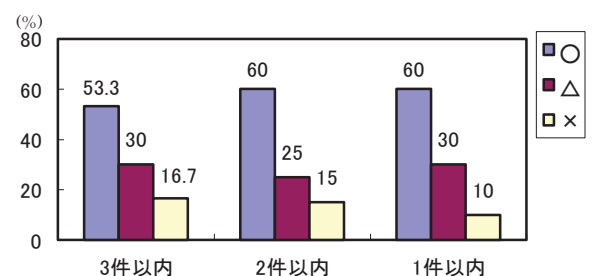


図9 入力と同じ提案をした場合を省いた評価  
 (料理概念ベースを用いた)

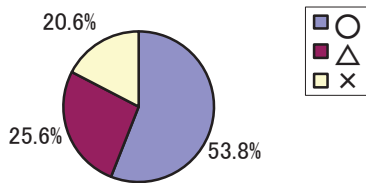


図 10 入力と同じ提案をした場合を省いた評価  
 (料理概念ベースを用いない)

次に料理概念ベースを用いない提案システムでは失敗したが、料理概念ベースを用いた提案システムで成功した入力例を表 9 に挙げて考察を行う。

表 9 入力例

好きな食べ物	シチュー	ポテトサラダ	チーズ
嫌いな食べ物	きゅうり	人参	甘いもの

表 9 のような入力があったとき、料理概念ベースを用いた提案システムではカレーを提案した結果○であった。しかし料理概念ベースを用いない提案システムではカレーが提案されていなかった。これは嫌いな食べ物に人参があったため、人参を含んでいる料理は省かれてしまうからである。評価者によっては、嫌いな材料を使用しているからと言ってその料理が嫌いな料理であるとは限らないと言える。料理概念ベースを用いた提案システムでは「シチュー」と「カレー」の料理関連度の値の方が「人参」と「カレー」の料理関連度の値よりも大きかったことで「カレー」を提案することができたと考えられる。

次に料理概念ベースを用いた提案システムの失敗例を図 11 に挙げて考察を行う。

入力

好きな食べ物	甘いもの	チョコレート	チーズ
嫌いな食べ物	トマト	豚肉	鶏レバー



出力

提案する料理はすき焼きです。

図 11 料理概念ベースを用いた提案システムの失敗例

図 11 の例は評価で×になった失敗例である。好きな食べ物が甘いもの、チョコレートなのでパフェやケーキなどの提案が予想されたが、実際はすき焼きが提案された。この原因として、6.1.1 節の処理で「甘いもの」から調味料「砂糖」が取得される。そして概念「すき焼き」の属性に砂糖が含まれていることにより、「砂糖」と「すき焼き」の関連度の値が高くなったと考えられる。調味料に砂糖が使われていても、甘い料理とは限らないと言える。この問題の解決策として、本稿で構築した料理概念ベースには材料や調

味料の分量を考慮していないので、それらの分量を属性に格納するか、属性の重みに反映することで改善されると考えられる。

最後に料理概念ベースの属性評価を目視で行った。属性の評価として、概念とそれぞれの属性を無造作に 20 セット取り出し、その概念に対して属性が正しいか否かを判断し、正解率の割合を精度とする。表 10 にテストデータの一部を示す。

表 10 料理概念ベースのサンプル概念とその属性

概念	属性
カレー	鶏肉, カレー粉, 塩コショウ, 玉葱, 人参, …
リゾット	米, 生クリーム, 玉葱, パルメザンチーズ, …
⋮	⋮

目視評価の結果、64%の精度を得ることができた。なお、料理概念ベースの全概念数は 3263 語で平均属性数は 21 語である。

また評価のために構築した、料理概念ベースを用いた料理提案システムでは出力として料理知識ベースのカテゴリの項目から料理を提案したが、出力として料理知識ベースの料理名を提案することでより詳細な料理の提案を行うことができる。

## 6. おわりに

本稿では料理同士の類似度を定量的に評価するために料理概念ベースの構築と料理関連度計算方式の提案を行った。また料理概念ベースを組み込んだ料理提案システムを構築し、ユーザの嗜好情報に合った料理の提案を行った。その結果、精度向上などの課題はまだ残されているが、料理概念ベースを用いた料理提案システム全体として 60%の精度を得ることが出来た。

**謝辞** 本稿の一部は、科学研究費補助（若手研究(B)24700215）の補助を受けて行った。

## 参考文献

- 1) 渡部広一, 奥村紀之, 河岡司, “概念の意味属性と共起情報を用いた関連度計算方式”, 自然言語処理, Vol.13, No.1, pp.53-74, 2006.
- 2) 奥村紀之, 土屋誠司, 渡部広一, 河岡司, “概念間の関連度計算のための大規模概念ベースの構築”, 自然言語処理, Vol.14, No.5, pp.41-64, 2007.
- 3) ChaSen -- 形態素解析器, <http://chasen-legacy.sourceforge.jp/>, 奈良先端科学技術大学院大学情報科学研究科自然言語処理学講座(松本研究室), 2014/1/17
- 4) 徳永健伸(編), “情報検索と言語処理”, 東京大学出版会, 1999.
- 5) NTT コミュニケーション科学研究所監修, “日本語語彙体系”, 岩波書店, 1997.
- 6) “E・レシピ”, <http://recipe.woman.excite.co.jp/>, 2014/1/17