

エンティティの構成要素の安定性を考慮したインタラクティブな Web エンティティ検索とそのレシピ推薦への応用

佃 洗撰[†] 中村 聡史[‡] 田中 克己[‡]

京都大学工学部情報学科[†] 京都大学大学院情報学研究科社会情報学専攻[‡]

1. はじめに

近年, Web 上で提供される情報やサービスは飛躍的に増加している. ユーザはそれらの情報について記述されたページを検索及び閲覧することにより, 京都の桜の観光に関する情報や料理のレシピなどのような多様な情報を得ることができる.

しかし, ユーザがあるページを閲覧した場合, ユーザはそのページに記述されている情報しか得ることができないため, 「もう少し記述を追加した情報が欲しい」または「ここから何か 1 つを省きたい」という意図をユーザがもつことが多々ある. 例えば, あるおでんのレシピを見ているユーザが, 物足りなさを感じてあと 1 つ食材を追加したいと考えたり, ある京都の桜の観光コースに関するページを見ているユーザが, 全ての観光地に行く時間はないのであと 1 箇所観光地を削除したいと考えるといった状況がある.

現時点で, 上記のような検索を可能とするサービスは存在しない. また, 一般的な検索エンジンを利用して上記のような検索を行うのは非常に困難である. そこで本研究では, こうしたある情報に対する, 何らかの要素の追加および削除に注目し, 特にレシピに対する食材の追加・削除を実現する.

2. エンティティ

ユーザが「おでん」のレシピについて調べているとき, 「京風おでん」や「関東風おでん」, 「洋風おでん」などのように, 様々なバリエーションがある. このそれぞれは大根や卵, ちくわぶや海老芋, ロールキャベツなど各種の具材の組み合わせにより変化するものである. そして, それぞれの具材の組み合わせは「おでん」という同じカテゴリ内のものである.

このように, ユーザが最終的に辿り着こうとしているある対象は, あるカテゴリに属しており, 海老芋なら京風おでん, ロールキャベツなら洋風おでんなどのように, その対象を構成する要素の組み合わせにより, その特性を変えるものである. 我々は, このように各

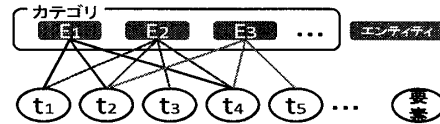


図 1: エンティティのモデル

要素の組み合わせにより形作られるものをエンティティとよぶ. 図 1 はこれをモデル化したものである.

本研究では, エンティティをその構成要素が土台とその上の層を構築するような, 2 層構造を成すものとして考える. 土台を構成する要素は, あるカテゴリに含まれるほぼ全てのエンティティに含まれる重要な要素であり, 上層を構成する要素は, いくつかのエンティティにのみ含まれる修飾的な要素である. 例えば, 「おでん」というカテゴリには, 「京風おでん」や「関東風おでん」, 「洋風おでん」など様々な種類のおでんが存在するが, どのような種類のおでんをつくる場合でも, 「卵」や「じゃがいも」といった食材は欠かせず, 土台となる食材である.

さらに, オブジェクトの構造を求めることで, 構造の安定度を考えることが可能になる. おでんの例では, 特異な具材を含まない一般的なおでんほど安定した構造であり, ロールキャベツなど一般的に使わない食材を含む洋風おでんは不安定な構造になる. 我々は, ユーザが選択したレシピの構造の安定度や, 食材を追加または削除した後の安定度を数値化することで, ユーザが目的のエンティティに辿りつくことを支援する. これにより, ユーザは今見ているレシピを典型的なレシピに近付けることや, 一風変わったレシピに近付けることが可能となる.

3. システムの概要

本研究では, レシピデータとしてインターネット上から得られたレシピについて, 各レシピのカテゴリや使われている食材等をデータベースに格納して用いる. システムの処理の流れは以下の通りである.

- (1) ユーザがレシピを選択すると, システムはそのレシピが含まれるカテゴリ内の全レシピ $E = \{e_1, e_2, \dots, e_n\}$ および全要素 $F = \{t_1, t_2, \dots, t_m\}$ を抽出. なお, n はカテゴリ内レシピの数, m は要素の数である.

Interactive Web Entity Search Based on Stability of Entity's Components and Its Application to Recipe Recommendation
[†]School of Informatics and Mathematical Science Faculty of Engineering, Kyoto University
[‡]Department of Social Informatics, Graduate School of Informatics, Kyoto University

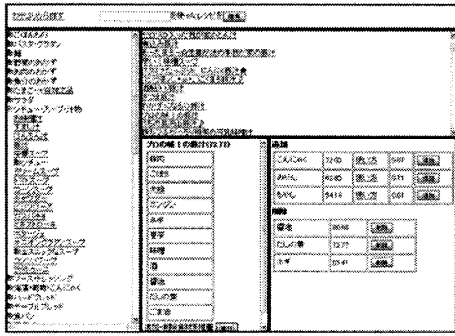


図 2: 実装したプロトタイプ

- (2) T の全ての t について出現頻度 $TF(t)$ を求め、 TF 値が閾値 α を超えるものを土台候補集合 S に分類。また、 B を土台に含まれる食材の集合とする。
 - ① S から、カテゴリ内での最も多くのレシピで使われている食材 s_i を求め、 s_i を S から B へ移動。
 - ② S の食材の中で、 B に含まれる食材群と最も共起する食材 s_j と、 B に含まれる食材との共起度 c を求める。
 - ③ c が閾値 β 未満であれば終了する。
 - ④ c が閾値 β 以上のとき、食材 s_j を土台となる食材として、 S から B へと移動し、手続き②に戻る。
- (3) カテゴリ内のレシピで使われる最も典型的な食材群を、(2)と同様にして求める。
- (4) ユーザが選択したレシピで使われている食材群と最も典型的な食材群との差異を元に、ユーザが選択したレシピの安定度を求める。
- (5) 追加・削除の候補の食材群に対して、各食材を追加・削除した場合のレシピの安定度を求め、その値が閾値以上となる食材の中から、安定度が最も高くなる食材と最も低くなる食材、安定度がその中間程度になる食材の 3 つをユーザに提示する。
- (6) ユーザが追加または削除する食材を選択すると、システムはユーザが選択したレシピで使われている食材群に対してその食材の追加・削除を行い、ユーザはさらに食材の追加及び削除を実行できる。

4. 評価

提案手法の有用性を評価するため、我々はシステムのプロトタイプを実装し、実験を行った(図 2)。「豚汁」と「和風サラダ」のカテゴリ内のあるレシピに対する実験結果は表 1 のようになった。表の中の安定度は、最も典型的なレシピの安定度を 100 とした時の、その時々々のレシピの安定度を表す。「豚汁」のカテゴリでは追加・削除共に、システムは安定度が大きく異

表 1 システムの出力

カテゴリ「豚汁」

追加		削除	
食材	安定度	食材	安定度
こんにやく	72.03	醤油	80.66
みりん	62.85	だし汁	73.77
もやし	54.14	ねぎ	65.41

レシピ選択時の安定度は 73.77

カテゴリ「和風サラダ」

追加		削除	
食材	安定度	食材	安定度
酢	18.90	白ゴマ	31.13
ニンジン	18.75	ツナ缶	22.88
水菜	18.38	醤油	17.53

レシピ選択時の安定度は 19.21

なる食材を提示した。追加食材の中で、元のレシピより安定度が高くなる食材が無かったのは、このレシピが典型的な食材を全て使っていたためである。また、通常「豚汁」に欠かせない「だし汁」が削除の候補として提示されたのは、使うことが自明な食材はレシピの食材リストに記載されていない場合もあり、土台となる食材に含まれなかったためであると考えられる。

「和風サラダ」は、カテゴリが広いので、多様な食材が使われる。そのため、土台となる食材や典型的な食材が少なく、どのレシピの安定度も低くなり、追加する食材の違いによる安定度の変化の差は殆ど見られなかった。精度の向上のためには、レシピの作り方から食材ごとの役割を分析する^[1]ことなどが考えられる。

5. まとめ

本論文では、あるカテゴリに属するエンティティに対する要素の追加及び削除を支援するための手法を提案した。提案手法では、エンティティの構造の安定度を定義し、要素の追加及び削除後の安定度の変化に注目した推薦を行った。今後は構造や安定度の求め方の改善やユーザ評価を行うこと考えている。

6. 謝辞

本研究の一部は、グローバル COE 拠点形成プログラム「知識循環社会のための情報学教育研究拠点」、計画研究「情報爆発時代に対応するコンテンツ融合と操作環境融合に関する研究」(研究代表者: 田中克己, A01-00-02, 課題番号 18049041)によるものです。ここに記して謝意を表すものとします。

[文献]

- [1] 志土地由香, 高橋友和, 井手一郎, 村瀬洋: “調理レシピテキストからの代替素材の発見” 人工知能学会第 22 回全国大会論文集, 1B1-2, pp. 1-4