

間接相関ルールを用いた代替食材の抽出

川口 美香

要旨

SNSの発展と健康志向の高まりにより、利用者によるレシピ投稿サイトが注目を集めている。料理レシピのコミュニティウェブサイトであるクックパッドもまたその一つである。クックパッドは、簡単に自分の料理ホームページを持つことができることや、食材を使った検索ができるなどの特徴を持ち、現在、アクセス数・登録レシピ数ともに日本一のレシピサイトである。

本研究では、レシピ投稿サイトからの知識発見の一つとして、代替可能食材の発見を目的とする。大量なレシピデータからユーザーが好みのレシピを見つけたとしても、アレルギーへの対応の他、より安価で健康的な食材を利用したい、より簡単に調理を行いたいなど代替可能食材を探す必要がある。これらの多様なニーズに応える一般的な枠組みとして、本研究では、代替可能食材の発見に間接相関ルールを適用する。間接相関ルールとは、食材集合で表現される条件のもとで、同じ役割を果たす、またはライバルの関係にあるアイテム（食材）の対を表すものであり、この対を代替可能食材と捉える。一般に抽出されるルールは膨大になるため、何らかの基準を用いてランキングを行う必要がある。本研究では、レシピや食材に付与された関連情報やオントロジーを利用し、料理手順や食材の種類、食材の組み合わせなど、種々の観点からランキングに利用される8つの評価関数を提案する。また、複数の評価関数を組み合わせて利用することも検討する。

提案した枠組みを評価するため、クックパッドデータセットに含まれる「オムライス」レシピ4535件及び「ハンバーグ」レシピ14277件を対象に、評価実験を行った。実験では、各評価関数に基づくランキング結果を比較するとともに、利用者アンケートを行い、抽出されるルールやそのランキングの妥当性について検証を行った。