

D-8-11

## 大量のレシピからの料理に関する 類似性評価知識発見手法について

Mining Similarity Assessment Knowledge in Cooking

大平 峰子 大園 忠親 新谷 虎松

Mineko Ohira Tadachika Ozono Toramatsu Shintani

名古屋工業大学 知能情報システム学科

Dept. of Intelligence and Computer Science, Nagoya Institute of Technology

### 1. はじめに

本研究の目的は、料理に関する知識を多数のレシピから自動的に発見することである [1]。レシピには、料理に関する様々な知識が含まれているが、これらの知識を自動的に発見することは容易ではない。本論文では、料理間の類似性を評価するための知識（類似性評価知識）の発見手法について述べる。類似性評価知識は、事例ベース推論における類似性評価に利用可能である。事例ベース推論とは、過去の問題解決事例を現在の問題解決に利用する推論方式であり、現在の問題と過去に解決した問題の類似性を判断するための機構を必要とする。類似性評価知識を自動的に発見するための技術は、事例ベース推論機構の構築に大きく貢献する。

### 2. 料理における類似性評価知識の表現と利用

あるレシピにおいて、2つの材料が同時に出現することを、2つの材料が共起しているという。本研究では、レシピ中に現れる材料名（例：たまねぎ、牛バラ肉）や処理名（例：煮る、焼く）、そして料理名（例：ハンバーグ）の出現を対象とした共起関係に注目する。

知識発見の分野において、大量のデータ中に存在する共起関係を結合ルールとして抽出する研究がある [2]。結合ルールは、「 $A \Rightarrow B$ 」の形をしており、「 $A$ を満たすならば  $B$  が成り立つ」「 $A$ が生じる時、同時に  $B$ が生じる」というアイテムの共起関係を表現することができる。本研究では、アイテムは材料名や処理名、料理名を表す語である。例えば、料理における結合ルールは、

ハンバーグ  $\Rightarrow$  ミンチ, タマネギ

と表現でき、「ハンバーグ」ならば、「ミンチ」と「タマネギ」が含まれる」ということを意味する。

ある同じアイテム  $A$  に対して共起する異なるアイテムがあった場合、それらのアイテムは  $A$  が出現するという条件の元で類似性を持つと考えられる。本研究では、このような類似性を表現する知識を類似性評価知識とし、共通するアイテムまたはその集合と、共通要素に対して共起関係にある異なったアイテム（またはその集合）の組み合わせで表現する。類似性評価知識は、結合ルールを応用して、共通部分を条件部、相違部分を結論とするルールとして表現する。例えば、

にんじん, たまねぎ, じゃがいも  $\Rightarrow \text{sim}([\text{牛肉}], [\text{鶏肉}])$

という類似性評価知識は、「“にんじん”、“たまねぎ”、そして“じゃがいも”が同時に含まれている場合に“牛肉”と“鶏肉”は類似している」ことを意味する。結論部は、材料名と材料名の組など、少ない要素数に限定する。

この知識を用いた2つのレシピの類似性評価は次の通りである。まず、2つのレシピに出現するアイテムの中で共通部分を見つける。次に、それ以外の相違している部分のアイテムのうちできるだけ多くのアイテムが、類似性評価知識によって類似していると対応付けられる場合に、これらのレシピは類似していると評価することができる。

### 3. 類似性評価知識発見のためのアプリオリの改良

本研究では、類似性評価知識を結合ルールとして発見するために、アプリオリアルゴリズム（以下アプリオリ）を改良した。アプリオリは、大量のデータから効率良く結合ルールを抽出するためのアルゴリズムである。アプリオリでは、ルールのサポートと確信度によって、ルールの有効性を判断する。サポートとはルールの条件部と結論部を満たすデータ

の割合であり、確信度とはルールの条件部を満たすデータの中でさらに結論部も満たすデータの割合である。アプリオリは、次の2ステップでルールを発見する。(1) 対象とするデータに対してサポートが閾値以上のアイテムの集合を求める。(2) 求められたアイテム集合を条件部と結論部に分割し閾値以上の確信度を持つルールを列挙する。料理に関する類似性評価知識の発見にアプリオリを適用するためには、以下のような3つの課題がある。課題(1)は生成されるアイテム集合の冗長性を減らす手法の確立である。アプリオリでは、サポートが閾値を超えるアイテム集合すべてをルール生成の対象とする。しかし、あるアイテム集合の異なる部分集合に対して共通点と相違点を見つけても、類似性を評価する知識とはなり得ない。課題(2)は、サポートと確信度に代わるルールの妥当性評価手法の確立である。結合ルールにおいて、条件部と結論部の間に相関関係がない場合にも、確信度が高い値を示してしまう可能性がある。課題(3)は、アイテム集合から類似性評価知識を生成するための手法の確立である。

これらの課題を解決するための、アプリオリの改良方法を示す。課題(1)に対しては、対象とするアイテム集合を、他のアイテム集合と包含関係にならないアイテム集合だけに限定することで解決した [1]。課題(2)に対しては、ルールの評価に統計的に相関関係を評価するカイ2乗検定を取り入れることで解決した。カイ2乗検定を用いたルールの評価は、ルールの条件部と結論部の相関関係を評価することで達成される [3]。課題(3)に対しては、アイテム集合からの知識の生成方法として以下の方法を提案する。まず、互いに包含関係のないアイテム集合に対して、共通点を持つアイテム集合の組を見つける。次に、それらのアイテム集合間で相違しているアイテムの中から、材料名同士、料理名同士などのアイテムのペアを選択する。そして、共通部分を条件部、ペアのアイテムを結論部として、条件部が共通しているルールをそれぞれ生成する。条件部には、共通部分全体ではなく、部分集合を用いる。部分集合を用いた条件部を持つルールのうち、確信度、カイ2乗値による評価の高いものを選択し、類似性評価知識として表現を整える。このように、条件部を結論部との相関関係の高いアイテム集合に限定することで、類似性評価知識の汎用的な利用が可能になる。

### 4. おわりに

本論文では、料理における類似性評価知識を結合ルールに基づく形式で発見するための手法について述べた。ここでは類似性評価知識をアプリオリを用いて発見するための3つの課題について述べ、それらの課題に対する解決方法を提案した。本手法を用いることにより、大量のレシピから事例ベース推論に適用可能な類似性評価知識を自動的に発見することができる。

### 参考文献

- [1] 大平峰子, 大園忠親, 新谷虎松: “料理における類似性評価のための知識の発見手法について”, 第14回人工知能学会全国大会論文集(2000)(掲載予定)。
- [2] R. Agrawal et al.: “Fast Discovery of Association Rule”, AKDDM, pp.307-328(1996).
- [3] 森下 真一, 中谷 明弘: “データマイニングの理論と実装”, コンピュータソフトウェア Vol.17 No.1, pp.59-72(2000).