

栄養バランスを考慮した料理レシピ検索

苅米 志帆乃

筑波大学図書館情報専門学群

藤井 敦

筑波大学大学院図書館情報メディア研究科

1 はじめに

毎日の食生活を豊かにするためには、「健康」と「食べる楽しみ」への配慮として、栄養バランスや種類の豊富さを考慮し、献立を決定する必要がある。そこで、様々な料理に関する知識や作り方が必要である。近年は、Webなどのオンライン情報でレシピを調べる方法がある。Webには料理の種類が豊富であるという利点がある。一方で、「レシピ にんじん」などで検索しても、ほとんどの場合は単品の料理しか検索されない。すなわち、複数の料理を組み合わせる献立を考える作業はユーザの役目である。そこで献立検索の機能があればユーザの負担が軽減される。

献立検索の例として、「しゃべる!DSお料理ナビ」¹というゲームソフトがあり、材料、サイドメニューの数、スープの有無などの検索条件によって、献立を自由に決めることができる。利田ら [1] は、ユーザに合わせて栄養的にバランスのとれた献立を作成するシステムについて研究した。しかし、種々の知識ベースが必要で拡張性に乏しい。

本研究は、料理どうしの関係を事前に定義しない拡張性が高い献立検索の手法を提案する。

2 料理レシピ検索

2.1 概要

本研究で提案するレシピ検索システムの構成を図1に示す。本システムは、料理名や材料などの検索条件を入力して「初期検索」を行う。次に出力されたレシピに基づいて、組合せの良さという観点で検索を行う。これは、栄養バランスの良い組合せを検索する献立検索である。具体的にはレシピどうしの組合せにおける関連度を計算し、関連度によってレシピに順位をつける。

図1は、オンライン処理とオフライン処理に大別される。オフライン処理では、Webからレシピを収集し、レシピから材料や様式などの情報を抽出し、レシピのデータを作成する。材料と食品群を対応付ける「食品群辞書」を用いて、食品群ごとに摂取量を計算し、食品群別摂取量のデータを作成する。材料の単位において、グラムに変換が必要な場合は、「グラム変換辞書」を利用する。辞書を用いる際は、辞書との柔軟な照合が必要となる。オンライン処理では、検索システムに検索条件を入

力し、条件に合致するレシピをデータベースから検索する。さらに、検索されたレシピに基づいて別のレシピとの関連度を計算し、献立検索を行う。対象とするレシピは、「キュービー3分クッキング」²のWebサイトから収集した。当レシピは項目のタグ付けが統一されていないため、料理名、材料、調理手順などの項目を手で抽出し整備した。整備したレシピは341件ある。

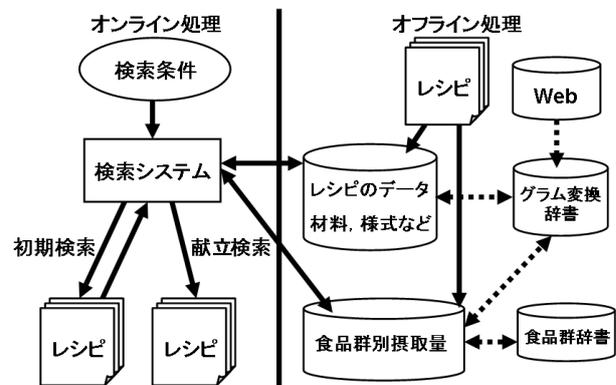


図 1: レシピ検索システムの構成

2.2 関連度の計算

栄養バランスを考慮して料理どうしの関連度を計算する。そこで、一日に摂るべき食品の分量に関する目安である「食品群別摂取量」を利用する。食品群別摂取量は、年齢や性別に応じて食品群ごとの摂取量が定められている。食品群とは、含まれる栄養素の種類によって食品を分けた集まりであり、基礎食品として表1に示す6つの食品群がある。食品群別摂取量で考慮される栄養素に、塩分は含まれていない。しかし、塩分は健康に関して重要な要素となるため、本研究では塩分を新たに第7群として追加する。また、カロリーが高い砂糖、マヨネーズ、ドレッシング以外の各種調味料（酢、しょうゆ、塩、ソース、こしょうなど）は食品群に分類されない。本研究では、分類処理の都合上、調味料・香辛料類を新たに第8群として追加する。

食品群別摂取量について具体的に説明する。例として20代、男性の場合を考える。20代、男性の1食で摂る

¹<http://www.nintendo.co.jp/ds/a4vj/>

²<http://www.ntv.co.jp/3min/>

表 1: 食品群の分類と属する食品の例

群	食品の例
第 1 群	魚介類(あじ, マグロなど) 肉類(牛肉, 鶏肉など) 豆類(大豆, 豆腐など) 卵類(鶏卵, うずら卵など)
第 2 群	乳製品(牛乳, チーズなど) 海藻類(のり, 昆布など)
第 3 群	緑黄色野菜(にんじん, トマトなど)
第 4 群	その他の野菜(キャベツ, きゅうりなど) きのこ類(しいたけ, しめしなど) 果物(りんご, みかんなど)
第 5 群	穀類(米, パンなど) イモ類(じゃがいも, さつまいもなど) 砂糖類(砂糖, はちみつなど) 菓子類(ケーキ, ビスケットなど)
第 6 群	油脂類(バター, ごま油)

表 3: 3つの料理に対する食品群別摂取量

食品群	肉じゃが	卵豆腐	チーズサラダ
第 1 群	50	75	0
第 2 群	0	0	10
第 3 群	0	0	18
第 4 群	31	0	16
第 5 群	150	0	0
第 6 群	0	0	3
第 7 群	2.6	1.2	1.3

$$\text{充足率 } i = \frac{|\text{規定量 } i - \text{過不足分 } i|}{\text{規定量 } i} \quad (1)$$

式(1)によってさらに、全食品群を横断して関連度を計算する。

$$\text{関連度} = \sum_{i=1}^7 \text{充足率}_i \quad (2)$$

べき摂取量を表 2 示す。

表 2: 20 代男性の食品群別摂取量

食品群	1 食で摂るべき摂取量
第 1 群	87
第 2 群	133
第 3 群	33
第 4 群	133
第 5 群	170
第 6 群	8
第 7 群	3

「肉じゃが」と「卵豆腐」を組合せた場合と「肉じゃが」と「チーズサラダ」を組合せた場合を比較する。肉じゃがには、牛肉(第1群)、玉ねぎ(第4群)、じゃがいも(第5群)が含まれている。卵豆腐には、卵(第1群)が含まれており、チーズサラダには、チーズ(第2群)、トマト(第3群)、レタス(第4群)、オリーブオイル(第6群)が含まれている。肉じゃが、卵豆腐、チーズサラダに対する食品群別摂取量を表 3 に示す。

表 2 と表 3 を比べると、肉じゃがと卵豆腐の組合せでは、第 1 群が超過し、第 2・3・4 群は摂取できない。すなわち、摂取できる食品群が偏っているため悪い組合せである。他方において、肉じゃがとチーズサラダの組合せでは、各食品群をまんべんなく摂取できるため、卵豆腐との組合せよりも良い組合せである。

食品群をまんべんなく摂取できているかどうかを判定するために、食品群ごとの充足率によって関連度を計算する。まず、複数の料理を組合せた場合の食品群別摂取量を合計し、ユーザの年齢と性別で規定されている量と比べて過不足分を計算する。過不足分に基づいて食品群 i の充足率 i を計算する。充足率 i は式 (1) で計算する。

以上の手法は、3つ以上の料理に対しても組合せの良さに関する関連度を計算することができる。

2.3 食品群辞書

食品群別摂取量を利用するために、レシピ中の材料名を食品群に分類する必要がある。そこで「食品群辞書」を作成する。これは、材料名とその食品群を記した辞書である。辞書は「カラーグラフ食品成分表」[2]を人手で電子化し作成した。作成した辞書の内容は、表 1 と同じである。

材料名は、食品群辞書を用いて照合できた材料名と同じ食品群に分類する。レシピから抽出した材料数は 502 あり、材料名を使って完全一致で照合できた材料名は 145 件しかなかった。そのため、異表記で照合できない材料名や辞書にない材料名への対処として、柔軟な照合が必要となる。そこで、以下の(A)から(F)を順番に実行し、分類できた段階で処理を終了する。

- 材料名をそのまま照合する。
- 材料名の読み仮名で照合する。
- 材料名の単位が「大さじ」、「小さじ」、「その他(適量, 少々など)」の材料は調味料に分類する。単位が「大さじ」、「小さじ」などは、ほとんどの場合、材料が調味料や香辛料である。
- 材料を形態素解析して、一部を照合する。
例えば、「りんごのすりおろし」を形態素解析すると、「りんご」、「の」、「すり」、「おろし」に分割される。これらのうち、「りんご」は第 4 群に属するため、「りんごのすりおろし」を第 4 群に分類する。
- N グラム検索と Cyclone³ で検索される関連語を使い、多数決によって分類する。N グラム検索は、構成文字が似ている材料名を検索する。例えば、「にん

³<http://cyclone.slis.tsukuba.ac.jp/>

じん」を構成するバイグラムは、「にん んじ んじ じん」である。構成文字が似ている材料名として、にんにく(にん んに にく)、ちりめんじゃこ(ちり りめ めん んじ じゃ やこ)などがある。

Cycloneの関連語検索は、検索窓に入力した語と関連がある言葉を検索する。Cycloneの検索窓に「にんじん」と入力した場合の出力画面を図2に示す。「にんじん」と入力すると、ビタミンA、カロチン、バターなどの関連語が検索される。レシピの材料名に存在しない「ビタミンA」や「カロチン」は関連語から削除する。

検索された関連語の食品群を調べ、一番多い食品群に材料名を分類する。例えば、「焼き豆腐」の関連語は、こんにゃく(5群)、しょうが(4群)、高野豆腐(1群)、木綿豆腐(1群)であるため、「焼き豆腐」を1群に分類する。



図2: 「にんじん」と入力した場合のCycloneの検索結果

2.4 グラム変換辞書

食品群別摂取量の単位はグラムである。しかし、レシピには「にんじん 2本」のようにグラム以外の単位で表記されている材料がある。そのため、レシピにある材料の単位をグラムに変換する必要がある。変換には、材料名と分量の単位、それに対するグラム表記を定義した「グラム変換辞書」を用いる。図4にグラム変換辞書の抜粋を示す。以下の方法でグラム変換辞書を作成した。

- 「カラーグラフ食品成分表」を手で電子化し作成した。辞書の項目数は95である。
- GoogleとYahoo!を用いて、材料名などでWebページを検索し、検索されたページから分量に関するパターンによって抽出された材料名と分量の単位、それに対するグラム表記の組合せを使って辞書を作成した。辞書の項目数は、Googleから50件、Yahoo!から67件追加した。例えば、「にんじん 1本g」という質問を用いて、Webページを検索する。検索結果を図3に示す。次に、括弧表現に関するパターンとの照合によって、数値を自動的に抽出する。図3から、「にんじん 1本(200g)」というパターンを抽出して、辞書に「にんじん 本 200」を辞書に登録する。

レシピから抽出した材料名とその分量の単位の組合せは743である。このうちグラム表記されている組合せの数は246あり、これらはグラムに変換する必要がない。また、分量が少々や適量で表記されている組合せの数が



図3: 材料名を用いてWebページを検索した結果

表4: グラム変換辞書の抜粋

材料名	単位	グラム表記(g)
酒	小さじ	5
酢	大さじ	15
えび	尾	350
かぶ	個	106
にんじん	本	200

83あり、これらはグラムに変換できない。そのため、グラム変換の必要がある残りの414について処理を行った。材料名を使って完全一致で照合できた材料名は112しかなかった。そのため、食品群辞書同様、異表記で照合できない材料名や辞書にない材料名への対処として、柔軟な照合が必要となる。そこで、食品群辞書の照合に用いた(A), (B), (D), (E)を順番に実行する。

3 評価

3.1 食品群辞書との柔軟な照合

2.3節で説明した(A)~(F)を順番に実行した場合に照合できた材料数、再現率、精度を表5に示す。処理が進むにつれ、精度は下がる。しかし、照合できる材料数は増え、再現率は79.1%になった。結果として全体の約80%が正しい食品群に分類できた。

表5: 食品群に分類する手法の評価

照合方法	材料数	再現率(%)	精度(%)
(A)	145	28.9	100
(B)	162	32.3	99.4
(C)	286	57.0	95.3
(D)	350	69.7	89.1
(E)	389	77.5	82.4
(F)	397	79.1	79.1

3.2 グラム変換辞書との柔軟な照合

2.4節で説明した辞書の作成方法(a)と(b)を用いて、2.3節で説明した(A), (B), (D), (E)を順番に実行した場合の再現率と精度を表6に示す。再現率は、グラム変換が必要な組合せに対して、正しくグラムに変

換できた割合である。また、精度はグラムに変換した組合せに対して、正しくグラムに変換できた割合である。グラムに変換する材料名と照合できた材料名が同じである場合は正解とし、異なる材料名を使ってグラムに変換した場合は誤りとする。例えば、「にんじん 本」を「にんじん 本 (200g)」ではなく、「バナナ 本 (150g)」などを使ってグラムに変換した場合は誤りである。

表 6: グラムに変換する手法の評価

照合方法	辞書	再現率 (%)	精度 (%)
(A)	(a)	8.45	100
	(b)	27.1	100
(B)	(a)	27.1	100
	(b)	27.3	100
(D)	(a)	30.7	94.8
	(b)	33.1	91.3
(E)	(a)	33.6	67.8
	(b)	34.5	59.8

処理が進むにつれ、精度は下がる。しかし、照合できる材料数は増え、再現率は 34.5% になった。結果として、グラムに変換する必要がない材料名を含めると全体の約 60% がグラム表記になった。(a) のみを用いて照合した場合は 52.4% で、(b) を用いることにより 63.5% まで増やすことができた。(b) は自動手法であり、簡単に辞書を更新することができるため、有効な手法である。

3.3 誤り分析

「食品群辞書」との柔軟な照合について、正しく照合できなかった材料名を手で分析した。正しく変換できなかった材料名とその単位の内訳を表 7 に示す。表 7 に基づいて誤りの原因について、考察する。

表 7: 食品群辞書と正しく照合できなかった材料の例

照合方法	材料名の例
(B)	酒
(C)	サラダ油, 干しエビなど
(D)	豆もやし, 米焼酎など
(E)	レーズン, ゆでたけのこなど

(B) では、「酒 (サケ)」と「鮭 (サケ)」のような同音異義語が存在する材料名が、正しく分類されなかった。分類精度を高くするためには、同音異義語の材料名リストを作成し、対象の材料名が当該リストに登録されている場合には、ユーザに確認する必要がある。

(C) は、単位が「大きじ」、「小さじ」、「適量」、「少々」である場合は、調味料に分類する。この 4 つの単位のうち、「適量」となっている単位の材料名は正しく分類されなかった。「適量」という単位だけで調味料に分類するべきかどうかは検討が必要である。

(D) では、形態素解析して、2 つ以上の単語が異なる食品群中の項目と照合された場合に誤りがあった。本手法では、最初に照合できた食品群に分類するため、「豆もやし」や「米焼酎」などは正しく分類されなかった。例えば、「豆もやし」は第 4 群に分類しなければならない。しかし、形態素解析すると、豆 (第 1 群)、もやし (第 4 群) となり、最初に照合された食品群に分類してしまうため、第 1 群に分類した。このような材料名は、あらかじめ辞書に登録しておく必要がある。

(E) では、料理などで一緒に使われる材料名が関連語として多く検索された。例えば、レーズンの関連語は、バター (第 6 群)、ベーキングパウダー (第 8 群)、生クリーム (第 2 群)、チーズ (第 2 群) などが検索されたために、分類誤りが生じた。

「グラム変換辞書」との柔軟な照合について、正しく照合できなかった材料名を手で分析した。正しく変換できなかった材料名とその単位の内訳を表 8 に示す。表 8 に基づいて誤りの原因について、考察する。

表 8: グラム変換辞書と正しく照合できなかった材料の例

照合方法	材料名の例
(D)	「米焼酎 カップ」、「うずらの卵 個」など
(E)	「練り辛子 小さじ」、「生クリーム カップ」など

(D) では、形態素解析して、2 つ以上の単語が異なる食品群中の項目と照合された場合に誤りがあった。例えば、「米焼酎 1 カップ」である。「米焼酎」を形態素解析すると、「米」と「焼酎」に分割され、材料名「米」で照合し、グラム変換を行う。そのため、辞書にある「米 1 カップ 160g」より、米焼酎の 1 カップが 160g と変換されてしまった。

(E) では、料理などで一緒に使われる材料名などが関連語として多く検索されたために誤りが起こった。例えば、「練り辛子 小さじ」は、関連語である「しょうゆ」としてグラムに変換してしまった。

4 おわりに

今後の課題として、対象とするレシピの件数を増やし、さらなる評価を行うことがある。また、レシピデータを Web などから自動で抽出することで、拡張性を高めることである。献立検索を行った場合、料理の種類が偏っている組合せや、各料理の種類が別々である組合せなどは、良い組合せではない。栄養バランスだけでなく、料理の種類や様式などを考慮して検索することも今後の課題である。

参考文献

- [1] 利田いずみ, 加久間勝. コンピュータを用いた料理献立システムー献立作成用エキスパートシステムー. 情報処理学会研究報告, No. 91-CH-10, pp. 1-8, 1991.
- [2] 実教出版出版部. カラーグラフ食品成分表. 実教出版.