

D-8-24

## オークション支援システムにおける 概念データベースを用いた財のキーワード抽出

Concept-based Keyword Extraction  
on Auction Support Systems

中野 昌弘 福田 直樹 新谷 虎松

Masahiro Nakano Naoki Fukuta 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

インターネットオークションは急成長している分野であり、その市場規模は数年のうちに40倍をこえるという予測もある。これにより出品される財の数が今以上に増加することが予想される。現在のオークションサイトはカテゴリによる分類を行い、関連する財のみを提示するようにしている。1つのカテゴリに1000件以上の財が存在することがよく起こる。この中から希望に最も合致する財を探すことを人手で行うには多大な労力を必要とする。キーワードマッチングによる検索も行うことができるが、これにはいくつかの問題がある。1) ユーザの希望する財と関連の高いキーワードがあること。2) 希望する財のキーワードが、省略や造語により様々な書き方をされる。1)は関連の高い言葉があれば、キーワードマッチングが有効となるということである。このような問題のため、キーワード検索であってもオークションから希望する財を探すことが難しい。

我々が提案するシステムは、あらかじめ作成した概念データベースの知識を使い、キーワードの正規化と財の分類を行う。分類した財、数十個を一覧表でユーザーに提示する。項目ごとのソートが可能であり、ユーザーはより自分の興味に沿った財から調べることができる。

### 2. システムのインターフェース

出品者によって書かれた財の情報は初め、文章で書かれている。これからキーワードのタイプと、そのキーワードを探し出す。一覧表ではこの2つは分類を行うための項目の種類と、その項目に対する財の情報になる。一覧表に表示される項目はユーザーが選択でき、必要な情報のみを表示することができる。ソートだけでなく、選択した財を一覧表から削除したり、特定の値を持つものを残すこともできる。

### 3. キーワードへの変換

テキストから、一覧表に利用するキーワードを得るためには、表現の揺れを吸収して正規化することが最大の問題となる。財の記述には、そのカテゴリに対するドメイン知識を想定した略語なども現れる。

キーワードを得るために知識を得る必要がある。この知識は1つのキーワードに対し、その項目名と、あるキーワードが現れたと判定するルールが記述されている。幾つかあるルールのうち1つでも完全にマッチすれば、そのキーワードが含まれているとする。ルール1つは単語の集まりで、

$$\alpha_1, \alpha_2, \dots, \alpha_n$$

というように書かれ、順番通りに単語が現れるかを調べる。

図1にテキストからキーワードを決定する処理手順を示す。初めに、1) 財のテキスト情報を、特定のカテゴリに特

化せずに単語に切り分ける。細かく切り分け過ぎたり、されなかつたりすることが起こる。2) 概念データベースの知識を使い、1つの単語、またはいくつかの単語を組み合わせてキーワードのルールとマッチングを行う。3) 推測されたキーワードを使い、財がどのようなものであるか同定を行う。4) 同定結果により、優先的に調べるキーワードを決定する。この結果は、単語の正規化でのルール検索に使われる。5) 3, 4を繰り返し行う。全ての単語を調べ、特定されるキーワードが変わらなくなるまで続ける。6) 得られたキーワードを結果として出力する。

この処理で、ルールとマッチングに失敗した単語が、単語への分割が不十分なために発生することがある。これを解決するため、失敗した時に特定の言葉が単語中に現れないか調べ、可能なら分割する。逆に分割され過ぎた場合は、ルールを長くすることで対応する。

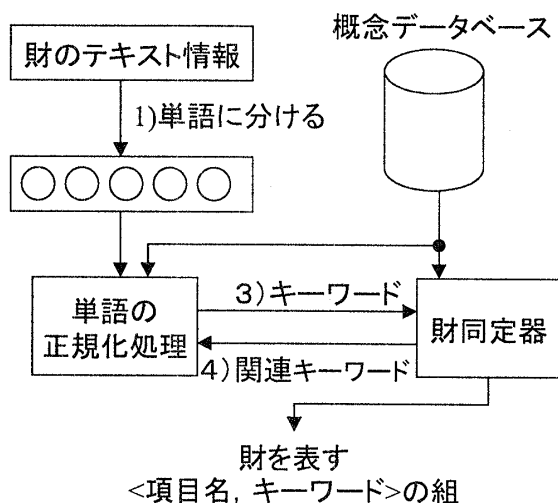


図1 キーワードの抽出過程

### 4. まとめ

キーワードマッチングによる財の検索の問題点を解決するため、ユーザーが、より柔軟な検索が行えるシステムを作成した。単語をキーワードへ正規化することで、キーワードマッチングでは漏れてしまう財も発見することができる。ただこれはルールの精度に依存しており、ルールを人手で作成するにはシステムと財、両方の知識が必要である。今後は、ルールの作成を自動化、半自動化することが課題にあげられる。

**参考文献** [廣田 99] 廣田啓一, 佐々木裕, 加藤恒昭. オントロジー主導による情報抽出. 人工知能学会誌 Vol. 14, No. 6, pp. 1010-1018, 1997