

平成 22 年 6 月 16 日現在

研究種目：若手研究（B）  
 研究期間：2008 ～ 2009  
 課題番号：20700192  
 研究課題名（和文） オブジェクト指向ラフ集合による構造と内容の相互作用を用いた嗜好抽出  
 研究課題名（英文） A study on preference extraction using interaction of structure and content based on object-oriented rough sets  
 研究代表者  
 工藤 康生（KUDO YASUO）  
 室蘭工業大学・大学院工学研究科・助教  
 研究者番号：90360966

## 研究成果の概要（和文）：

本研究では半構造データとして表現された対象の構造と内容の相互作用に着目し、対象に関するユーザの嗜好を、研究代表者が提案したオブジェクト指向ラフ集合を用いて抽出する手法について研究を行った。その結果、オブジェクト指向ラフ集合における構造と内容の相互作用の取り扱いについて、理論的基礎をある程度構築することができた。また、嗜好抽出手法の効率化および抽出結果の評価手法についても研究を行った。

## 研究成果の概要（英文）：

In this study, we proposed extraction methods of users' preferences from semi-structured data using object-oriented rough set models. As the results of this study, we developed a part of theoretical basis for interactive treatments of structures and contents of given semi-structured data based on object-oriented rough set models. Moreover, we also proposed refined preference extraction methods and evaluation methods of extracted results.

## 交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008 年度	1,000,000	300,000	1,300,000
2009 年度	600,000	180,000	780,000
年度			
年度			
年度			
総計	1,600,000	480,000	2,080,000

## 研究分野：総合領域

科研費の分科・細目：情報学・感性情報学・ソフトコンピューティング

キーワード：ラフ集合，半構造データ，特徴抽出

## 1. 研究開始当初の背景

近年、ラフ集合を用いたアンケートデータ等からのユーザの嗜好抽出は、データマイニング技術の感性情報処理への応用として注目を集めている。嗜好の抽出に関する従来研

究では、対象（主に商品）に関する様々な性質と、その対象の好ましさと明示的な関連性を抽出することに重点が置かれている。しかし、対象に関する情報が XML 等の半構造データで表現される場合は、対象が持つ構造

が内容の評価に影響を与えること、逆に対象の内容は構造の評価に影響を与えることから、半構造データで表現される対象からの嗜好抽出では、ユーザが好むと思われる構造と内容の両方を抽出する必要がある。

研究代表者はラフ集合にオブジェクト指向の概念を導入したオブジェクト指向ラフ集合を提案し、従来のラフ集合では取り扱いが困難だった半構造データに対してラフ集合の手法を用いることができる枠組みの定式化を行ってきた。また、感性情報処理への応用として、オブジェクト指向ラフ集合による半構造データからのユーザの嗜好抽出に取り組んでいたが、構造と内容の相互作用は考慮されていなかった。

## 2. 研究の目的

本研究は、構造と内容との相互作用を利用し、構造的長と内容的長の特徴の両方を抽出することに焦点をあて、オブジェクト指向ラフ集合を用いて、半構造データとして表現された対象の集合からユーザの嗜好を抽出する手法を構築することを目的とする。

## 3. 研究の方法

本研究では、半構造データにおける構造と内容の相互作用に着目し、以下の3つの研究を行う。

(1)ユーザの嗜好に合致する内容的特徴から構造的特徴を抽出するモデルを構築する。研究代表者は従来研究で、半構造データの構造的特徴を用いて内容的特徴を抽出する手法を提案しているため、内容的特徴を踏まえて構造的特徴を抽出することができれば、オブジェクト指向ラフ集合において構造と内容の相互作用を取り扱うことが可能となる。

(2)オブジェクト指向ラフ集合における嗜好抽出手法の構築およびその改良を行う。これはオブジェクト指向ラフ集合での相対縮約および決定規則の抽出に基づくため、これらの手法およびその改良を行う。また、その基礎として通常のラフ集合における相対縮約の抽出手法についても検討を行う。

(3)オブジェクト指向ラフ集合による嗜好抽出の評価手法を構築する。上記の通り、オブジェクト指向ラフ集合による嗜好抽出は相対縮約および決定規則の抽出に基づくことから、嗜好抽出の評価手法の基礎として、特に相対縮約の評価手法を構築しその有用性を検証する。

## 4. 研究成果

### (1)研究の主な成果

①オブジェクト指向ラフ集合における構造と内容の相互作用について、半構造データの

内容的特徴から、ユーザの嗜好に合致する可能性がある構造的特徴を推測し抽出するモデルの構築を試みた。具体的には、オブジェクト指向ラフ集合において以下の形式で構造と内容の相互作用を扱うモデルの構築を試みた。

- 構造から内容への作用：オブジェクト指向ラフ集合において対象間の構造的つながりを規定するクラス構造を用いて、各クラスのオブジェクトとして記述される対象に関する、対象間の規則性を抽出する。
- 内容から構造への作用：ユーザの嗜好に合致する対象から、それらの対象に共通する構造的な規則性を抽出し、新たなクラスとしてオブジェクト指向ラフ集合モデルに追加する。

構造から内容への作用に関する取り扱いについては研究代表者の従来研究で考慮されているため、本研究では内容から構造への作用の取り扱いについて研究を行った(学会発表③)。まず、複数のオブジェクトを組み合わせた複合的なオブジェクトの構造が、その複合的なオブジェクトに対するユーザの評価に及ぼす影響として、以下に示す正の影響と負の影響について考察した。

- 正の影響：評価が良くないオブジェクトも含まれているにもかかわらず、これらを組み合わせて得られた複合的なオブジェクトがユーザから良い評価を得る場合。
- 負の影響：評価の良いオブジェクトのみを組み合わせたにもかかわらず、複合的なオブジェクトがユーザからの良い評価を得られない場合。

複合的なオブジェクトの構造が持つ正の影響については、その構造が持つ「組合せの良さ」が、複合的なオブジェクトに対するユーザの良い評価をもたらしていると考えられることができる。逆に、負の影響については、その構造が持つ「組合せの悪さ」が、複合的なオブジェクトの低い評価の要因と見なすことができる。

この考察に基づき、複合的なオブジェクトに対して上記の正の影響(または負の影響)をもたらしていると思われる構造を、オブジェクト指向ラフ集合モデルのクラス構造から抽出する手法を提案した。これにより、構造から内容への作用をオブジェクト指向ラフ集合において取り扱うための理論的基礎をある程度構築することができた。

②嗜好抽出手法の構築および改良として、与えられたオブジェクト指向ラフ集合モデルにおける、条件部の長さが極小かつ条件部の

記述内容ができるだけ詳細な半構造決定ルール (semi-structured minimal decision rule, 以下 SSMDR と略記) をすべて抽出するアルゴリズムを提案した (学会発表⑤). 半構造決定ルールは以下の形式の決定ルールである.

- 条件部: クラス  $C$  のオブジェクトでその特徴  $p_1, \dots, p_i$  が  $C$  のあるオブジェクト  $x$  と等しい.
- 結論部: 条件部を満たすオブジェクトが持つ別の特徴  $q_1, \dots, q_j$  も  $x$  と等しい.

SSMDR は条件部に現れる特徴の個数  $i$  が極小であり, かつ各特徴  $p_1, \dots, p_i$  ができるだけ詳細となる半構造決定ルールである. 提案したアルゴリズムは, 通常のラフ集合においてすべての極小決定ルールを抽出する手法である決定行列に基づいている.

なお, 本節①で述べた構造と内容の相互作用の定式化では, 正の影響 (または負の影響) を抽出するために SSMDR を用いるため, 本アルゴリズムを用いることにより, 構造と内容の相互作用の取り扱いの基礎部分が可能となった.

嗜好抽出に関連して, 通常のラフ集合における相対縮約の粒状推論への応用 (雑誌論文③, 学会発表④) および相対縮約のヒューリスティックな計算手法 (雑誌論文①, 学会発表②) についても研究を行った. 粒状推論への応用は, 抽出した相対縮約から得られた概念 (ある特徴を満たすデータの集合) を用いた論理的推論の定式化を意図している. また, 相対縮約のヒューリスティックな計算手法については, 一般的に, すべての相対縮約を求める問題 (およびすべての極小決定ルールを求める問題) は NP 困難であることが知られているため, ある程度以上の規模のデータに対して相対縮約および決定ルールの抽出を行うためには, 何らかのヒューリスティックな計算手法が必須となる. 通常のラフ集合に関して行った上記の研究をオブジェクト指向ラフ集合に対して適用することは, 今後の課題とする.

③嗜好抽出の評価手法の基礎として, 通常のラフ集合における相対縮約および決定ルールの評価手法を構築しその有用性を検証した (雑誌論文②, 学会発表①および⑥). 提案した一連の手法では, 相対縮約を用いたデータの分類は元のデータの分割を与えることに着目し, 分類対象のデータをより正確かつ粗く分類する相対縮約ほど高く評価する. また, 提案手法で高評価を得た相対縮約から得られた決定ルールは, 低評価の相対縮約から得られた決定ルールと比較して, 平均的に多くのデータに当てはまることが保証されるため, より一般的な決定ルールとなること

が示された. 同様のアイディアはオブジェクト指向ラフ集合に対しても適用可能であると考えている. 通常のラフ集合に関して行った上記の研究をオブジェクト指向ラフ集合に対して適用することは, 今後の課題とする.

(2) 得られた成果の国内外における位置づけおよびインパクト

オブジェクト指向ラフ集合モデルは研究代表者らが独自に提案したモデルであり, ラフ集合で半構造データを扱う研究自体が, 国学も含めて, 申請者らの研究以外はほとんど存在しない. また一般的に, データマイニングの感性情報処理への応用研究では, 既知のデータからユーザが好むと思われる内容を見出すことに重点が置かれている. これに対し本研究では, 構造と内容の相互作用を利用することで, 半構造データからユーザが好むと思われる内容を抽出することに留まらず, ユーザが好むと思われる組み合わせ方自体も抽出し, 抽出した新たな組み合わせ方を用いて既存のデータには存在しない内容を提案することも可能にする. 以上より, 本研究はラフ集合の理論的側面に関する研究とデータマイニングの感性情報処理への応用研究の両者に対して, 新たな知見をもたらすことが期待できる.

(3) 今後の展望

本研究の申請期間は終了したものの, 本研究の内容はまだ発展途上であるため残された課題が多く, 今後も研究を継続する予定である. 継続する主な内容として以下が考えられる.

- 構造から内容への作用をもたらす構造的特徴の抽出を改良・厳密化する. これにより, 構造と内容の相互作用を用いた嗜好抽出を行う際に, ユーザの嗜好をより詳細に抽出し, かつ嗜好の変化等にも動的に対応することが可能になると予想される. また, 構造と内容の相互作用を用いた嗜好抽出手法の有効性についても引き続き検討する.
- SSMDR のヒューリスティックな計算手法の構築および SSMDR を用いた粒状推論の定式化. オブジェクト指向ラフ集合は記述論理 (description logic) と親和性が高いと考えられるため, オブジェクト指向ラフ集合と記述論理の関連性についても考察を行う.
- オブジェクト指向ラフ集合における相対縮約の評価手法を構築する. 前述の通り, 研究代表者らが提案した, 通常のラフ集合における相対縮約の評価手法と同様のアイディアがオブジェクト指向ラフ集合にも適用できると思われる.

## 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 3 件)

- ① Yasuo Kudo and Tetsuya Murai, Heuristic Algorithm for Attribute Reduction Based on Classification Ability by Condition Attributes, Journal of Advanced Computational Intelligence and Intelligent Informatics, 査読有, Vol. 15, No. 1, to appear.
- ② Yasuo Kudo and Tetsuya Murai, An Evaluation Method of Relative Reducts Based on Roughness of Partitions, International Journal of Cognitive Informatics and Natural intelligence, 査読有, Vol. 4, No. 2, in press.
- ③ Yasuo Kudo, Tetsuya Murai, and Seiki Akama, A Granularity-Based Framework of Deduction, Induction and Abduction, International Journal of Approximate Reasoning, 査読有, Vol. 50, Issue 8, pp. 1215-1226, 2009.

[学会発表] (計 6 件)

- ① Yasuo Kudo, On a Criterion of Similarity between Partitions Based on Rough Set Theory, The 12th International Conference on Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, 2009年12月16日, デリー (インド).
- ② Yasuo Kudo, A Heuristic Algorithm for Attribute Reduction Based on Discernibility and Equivalence by Attributes, The 6th International Conference on Modeling Decisions for Artificial Intelligence, 2009年11月30日, 淡路島 (日本).
- ③ Yasuo Kudo, On an Extraction Method of Structural Characteristics in Object-Oriented Rough Set Models, 2009 IEEE International Conference on Granular Computing, 2009年8月17日, 南昌(中国).
- ④ Yasuo Kudo, Attribute Reduction as Calculation of Focus in Granular Reasoning, The 39th International Symposium on Multiple-Valued Logic, 2009年5月23日, 那覇 (日本).
- ⑤ Yasuo Kudo, Decision Matrices for Semi-Structured Decision Rules in Object-Oriented Rough Sets, International Workshop on Soft Computing for Knowledge Technology, 2008年12

月16日, ハノイ (ベトナム).

- ⑥ Yasuo Kudo, An Evaluation Method of Relative Reducts Based on Roughness of Partitions, The 6th International Conference on Rough Sets and Current Trend in Computing, 2008年10月25日, アクロン (米国).

## 6. 研究組織

(1) 研究代表者

工藤 康生 (KUDO YASUO)

室蘭工業大学・大学院工学研究科・助教

研究者番号：90360966