

顧客レビューテキスト解析に基づく
文書作成支援に関する研究

2016年3月

角田 孝昭

顧客レビューテキスト解析に基づく 文書作成支援に関する研究

角田 孝昭

システム情報工学研究科
筑波大学

2016年3月

目次

第 1 章	はじめに	1
1.1	研究の背景と目的	1
1.2	本論文の構成	5
第 2 章	関連研究	7
2.1	文書作成支援に関する研究	7
2.1.1	文書作成支援の分類	7
2.1.2	構想過程の支援	8
2.1.3	文書化過程の支援	10
2.1.4	校正・推敲過程の支援	11
2.2	顧客レビュー文書を用いた研究	12
2.2.1	評判分析研究の概観	12
2.2.2	評判分析の応用	13
2.3	本研究の位置付け	14
第 3 章	商品紹介文に相応しい評価視点の提示	15
3.1	研究課題の背景と概要	15
3.2	評価視点・評価視点グループ	16
3.2.1	評価視点	16
3.2.2	評価視点グループ	16
3.3	基盤技術	17
3.3.1	評価視点表現の評価視点グループへの対応付け	17
3.3.2	対数尤度比	21
3.4	提案手法	23
3.4.1	提案手法の枠組み	23
3.4.2	評価視点グループの定義	24
3.4.3	評価視点表現の評価視点グループへの対応付け	25
3.4.4	評価視点グループのランキング	25
3.5	評価実験	28
3.5.1	実験条件	28
3.5.2	実験結果	30
3.6	関連研究	34

3.7	まとめと今後の課題	35
第4章	対をなす二文書間における文対応関係の推定	37
4.1	研究課題の背景と概要	37
4.2	文種類・文対応	40
4.2.1	文対応	40
4.2.2	文種類	40
4.2.3	文種類と文対応データの分析	43
4.3	基盤技術	47
4.3.1	Conditional Random Fields	47
4.3.2	因子グラフ	51
4.3.3	Linear-chain Conditional Random Fields (Linear-chain CRF)	52
4.3.4	2D Conditional Random Fields (2D CRF)	53
4.4	提案手法	53
4.4.1	問題の定式化	54
4.4.2	Linear-chain CRF による文種類推定	54
4.4.3	Linear-chain CRF による文対応推定	55
4.4.4	2D CRF による文対応推定	55
4.4.5	文種類・文対応推定モデルの統合	57
4.5	評価実験	59
4.5.1	実験条件	59
4.5.2	実験結果と考察	61
4.6	関連研究	67
4.7	まとめと今後の課題	68
第5章	おわりに	69
付録A	レビュー文書・応答文書における文分割手法	71
参考文献		75
謝辞		87
公表済み論文一覧		89

目次

1.1	EC サイト「楽天トラベル」における検索画面の例	3
1.2	EC サイト「楽天トラベル」におけるレビュー投稿と応答の例	4
2.1	文書作成の流れ	7
2.2	文書作成支援の分類	8
3.1	提案手法の枠組み	24
3.2	Revise セットと各ランキング手法による出力結果の分布	32
4.1	レビュー文書 – 応答文書対における文対応の例	38
4.2	コサイン類似度の分布（文対応なし）	47
4.3	コサイン類似度の分布（文対応あり）	47
4.4	因子グラフの例	51
4.5	Linear-chain CRF の因子グラフ表現	52
4.6	2D CRF の因子グラフ表現	53
4.7	Linear-chain CRF を用いた文対応の推定手順	56
4.8	統合前のモデル	57
4.9	統合後のモデル	58
4.10	2D-CRF と combine 間において出力が変化した文対応に対応する文種類のエントロピー	63
4.11	Precision-Recall 曲線	64
4.12	平均コサイン類似度値-Recall 曲線	64
4.13	combine モデルによる推定例	65
4.14	combine モデルと 2D CRF モデルの推定例	66
A.1	句点のみで文分割を行った例	71
A.2	本章で示す手法により文分割を行った例	72

表目次

3.1	評価視点グループの一覧	26
3.2	ポジティブ文分類器の設定	29
3.3	評価視点表現の評価視点グループへの対応付けの性能	31
3.4	As-is セットでの設定における評価視点グループのランキングの性能	31
3.5	Revise セットでの設定における評価視点グループのランキングの性能	32
3.6	S-LLR による出力例	33
4.1	レビュー文書を構成する文の種類	41
4.2	応答文書を構成する文の種類	42
4.3	レビュー文での文種類の分布	44
4.4	応答文での文種類の分布	44
4.5	レビュー文種類の遷移確率例	44
4.6	応答文種類の遷移確率例	45
4.7	レビュー文ごとの対応数・対応存在率	45
4.8	応答文ごとの対応数・対応存在率	45
4.9	文対応が存在する文種類の組み合わせ	46
4.10	文対応の遷移確率（レビュー文基準）	46
4.11	文対応の遷移確率（応答文基準）	46
4.12	文対応推定性能	61
4.13	提案モデル combine におけるレビュー文種類ごとの文対応推定性能	62
4.14	提案モデル combine における応答文種類ごとの文対応推定性能	62
4.15	2D CRF と combine 間における出力の変化（正解: 文対応あり）	62
4.16	2D CRF と combine 間における出力の変化（正解: 文対応なし）	62

第 1 章

はじめに

1.1 研究の背景と目的

ワードプロセッサの発明は、人類の文書作成活動に対して革命的な影響をもたらした。ワードプロセッサは、従来の手書きによる文書作成やタイプライターと比較して数多くの利点を有している。具体的な例としては、文字・文・段落の挿入・削除・置換が容易に可能である編集性の高さや、文書を保存しておくことで改めて利用できる再利用性の高さなどがある。ワードプロセッサがソフトウェアの形でパーソナルコンピュータに搭載されるようになった今日では、コンピュータの計算・記憶能力を活かすことで、単なる文書作成機能にとどまらず、スペル誤りや文法誤りの検出・訂正をはじめとした人間の文書作成を支援する試みが行われている。本論文では、このような試みをコンピュータによる文書作成支援 (Computer-Aided Writing, Computer-Assisted Writing) と定義し、以下では単に文書作成支援と呼ぶ。

文書作成支援は、1971 年に Ralph E. Gorin によって開発されたとされるスペル誤り検出プログラム SPELL (Peterson, 1980) 以降、コンピュータの発達に伴って「性能の向上」と「支援機能の多様化」の 2 つの方向で進化を遂げた。性能の向上については、例えば英語におけるスペル誤り検出課題においては Golding & Roth (1999) が約 95% の accuracy の達成を報告しているなど、実用的な性能を持つ手法が数多く提案されている。一方、支援機能の多様化については、スペル誤りに対する訂正候補の提示や文法誤りの検出・訂正、更には文書を形にするまでの「どう書くべきか」の提案や「何を書くべきか」の提案へと拡大する方向への研究も進められている。その極点にあるのは、文や文書の自動生成、例えば物語 (ナラティブ) 生成 (Callaway & Lester, 2001) や応答発話生成 (Ritter et al., 2011; Shin et al., 2015; Sordoni et al., 2015) であろう。これらの研究が更に発展すれば、将来的にはコンピュータとの協調による文書作成、ひいてはあらゆる文書の半自動・全自動生成が可能になるかも知れない。

本研究では、以上に挙げた 2 つの方向のうち支援機能の多様化を推し進めるために、新たに電子商取引ウェブサイト (以下、EC サイトと略する) において商品販売者が作成する文書を対象して「何を書くべきか」の提案を行う課題を取り上げる。EC サイトを対象とする主な理由は、以下の 2 点である。

EC 市場規模の拡大：EC 購買活動は年々増加しており、経済産業省の調査によれば日本国内の 2014 年時点の消費者向け EC 市場規模は 12.8 兆円まで拡大した (経済産業省, 2014)。具体的な店舗数では、例えば Amazon.co.jp^{*1} は約 18 万店舗 (日本経済新聞, 2015)、楽天市場^{*2} は約 4 万店舗、楽天

^{*1} Amazon.co.jp: <https://www.amazon.co.jp/>

^{*2} 楽天市場: <https://www.rakuten.co.jp/>

トラベル^{*3}は約8万施設^{*4}もの販売事業者を擁するなど、ECはあらゆる事業者にとって重要な販売チャネルとなっている。

ECサイトにおけるテキストの重要性：ECは購入までのプロセスで実際の商品を手にとることが難しいという特性を持つことから、ウェブサイト上の様々なテキストが顧客の購買判断材料に用いられる。そのため、テキストの内容は購買率向上のための重要な役割を担っており、作成支援による効果が高い。

次に、ECサイトの文書のうち、本研究で対象とする文書について説明する。ECサイトに関連して販売者が作成する文・文書には、以下が挙げられる。

1. ウェブサイトの利用案内
(店舗紹介、規約、利用方法、決済方法・送付方法・納期等を含む、店舗の利用ガイド)
2. 各商品の紹介・説明文
3. 店舗や各商品カテゴリ・各個別商品などのキャッチコピー
4. 商品の販促文書(キャンペーン紹介、特集記事、ダイレクトメール等)
5. 顧客から寄せられた感想・評価・質問・要望・苦情への対応
(メール・ウェブフォーム・コールセンターでの問い合わせ、レビュー投稿への応答)

本研究では、この中でも特に顧客の購買判断材料として重要である「各商品の紹介・説明文」および「顧客から寄せられた感想・評価・質問・要望・苦情への対応」を対象とする。以下、それぞれの文書、及び応答文書と関連してレビュー文書について、実際に以降の実験データセットとしても使用するECサイト「楽天トラベル」を例に説明する。楽天トラベルはインターネットを介して全国のホテル・旅館・ペンションなどの宿泊施設予約を行えるサービスであり、この場合の商品は「宿泊サービス」となる。

商品の紹介・説明文(以下、商品紹介文と呼ぶ)とは、各商品の簡潔な紹介や誘い文句を含む、商品検索画面や商品一覧画面などで顧客に提示される文である。図1.1に、楽天トラベルにおいて商品紹介文を含む画面の例として、「温泉」をクエリ入力として商品(宿泊施設)を検索した画面を示す。画面中央の商品名下に書かれているテキストが、当該商品の紹介文である。ECサイトでは大量の商品が高速に閲覧されることから、商品紹介文は検索・一覧画面から商品の詳細・購入画面へと顧客を誘導させるための重要な要素である。更に、楽天トラベルをはじめとするモール型ECサイト^{*6}では、類似する商品との差別化も求められる。

顧客レビュー文書(以下、レビュー文書と呼ぶ)とは、実際に商品を購入・利用した顧客によって感想・評価・質問・要望・苦情などが書かれた文書である。楽天トラベルでは「クチコミ・お客様の声」がレビュー文書に相当する。図1.2は、楽天トラベルのある宿泊施設に対する顧客レビューを含む画面の例であり、左下のテキストがレビュー文書である。レビュー文書はその商品を実際に利用した顧客のみ^{*7}が知りうる情報が含まれていることから、様々な活用可能性を有する。顧客の視点で見ると、多く

^{*3} 楽天トラベル: <http://travel.rakuten.co.jp/>

^{*4} 楽天市場契約企業数: <https://www.rakuten.co.jp/com/inc/rc/info.html>。楽天市場と楽天トラベルの契約企業数は、2015年12月10日アクセス時に掲載されている、同年12月1日時点のデータ。

^{*5} 図1.1の各注釈は説明の都合上、著者が施したものである。なお、商品名(宿泊施設名)など、商品が特定できる情報には画像処理を施した。図1.2も同様である。

^{*6} 異なる販売事業者が、ECモールの運営事業者が提供するECプラットフォームを通じた出店を行うタイプのECサイト。

^{*7} なお、ECサイトによっては、実際にレビューの対象商品を購入していなくてもレビューを投稿可能な場合もある。楽天トラベルの場合、実際に楽天トラベル経由で予約・利用した場合のみ、利用施設のレビューを投稿できるシステムとなっている。

The screenshot shows the Rakuten Travel website interface. At the top, there are navigation links and a search bar. The main content area displays search results for hotels, including filters for area (全国) and type (指定なし). A list of hotels is shown with their names, ratings, and prices. A detailed view of a hotel listing is visible, featuring a '商品紹介文' (Product Introduction Text) label pointing to a link. The page also includes a sidebar with 'あなたにオススメの特集' (Special Features Recommended for You) and a '検索条件' (Search Conditions) section.

図 1.1: EC サイト「楽天トラベル」における検索画面の例⁵

の EC サイトではレビュー文書の投稿が誰でも閲覧可能となっているため、購買の判断材料として用いられる。実際に Chevalier & Mayzlin (2006) は Amazon.com⁸ 及び Barnes & Noble⁹ の、Hu et al. (2008) は Amazon.com の商品の売上とレビュー投稿を分析し、商品の売上傾向はレビュー投稿と関係があることを明らかにしている。特に楽天トラベルのように商品が無形のサービス財である場合には、販売者が提供する情報のみから接客などの品質を評価することは困難であるため、購買判断のための情報を増やすためによりレビュー投稿が重要視される (長島, 2009)。販売者の視点でも、顧客からの生の声は商品や商品広告手段などを反省・改善するための貴重な情報である。

レビュー投稿への応答 (以下、応答文書と呼ぶ) とは、顧客レビューに対する商品販売者からの応答文書である。楽天トラベルをはじめとする一部の EC サイトでは、レビュー文書に対する販売者の応答をサイト訪問者が閲覧できるようになっている (楽天トラベルの場合、応答するかは各販売者の任意である)。図 1.2 の右下のテキストが、左下のレビュー文書に対する応答文書である。応答文書にはレビュー文書で書かれた情報の補足、質問・要望への回答、苦情への謝罪・対応が含まれていることが多く、レビュー文書と同様に購買判断の材料に用いることができる。中でも、サービスの失敗に対する説明や謝罪などの顧客対応 (recovery) は顧客の印象改善に重要であり (Bitner et al., 1990)、特に対応の質が高い場合には最初にサービス失敗が無かった場合よりも顧客満足度が高くなる recovery paradox と呼ばれる現象が観察される場合があると言う調査もある (Hocutt et al., 2006)¹⁰。これとは逆に、応答文

⁸ Amazon.com: <https://www.amazon.com/>

⁹ Barnes & Noble: <http://www.barnesandnoble.com/>

¹⁰ 先行研究によっては recovery paradox の効果は小さい (Michel & Meuter, 2008)、又はない (Kau & Loh, 2006) としている調査もあるが、本研究では苦情を含むレビュー文書に対する適切な応答 (recovery) は recovery paradox の有無に関わらず

楽天トラベルトップ > 全国 > 兵庫県 > 神戸市内・有馬温泉 > クチコミ・感想・情報

★★★★★ 3.39 クチコミ・お客様の声(176件) この宿泊施設をお気に入り追加 メルマガ 幹事さん機能 友達にメール シェアする

施設紹介 プラン一覧 写真・動画(18) 地図・アクセス お客様の声(176) クーポン一覧

国内宿泊
 ANA 航空券+宿泊
 JAL 航空券+宿泊
 日帰り・デユース
 日付未定
 チェックイン
 2015/11/25
 チェックアウト
 2015/11/26

ご利用部屋数
 1 部屋
 ご利用人数
 1 部屋目:
 プラン検索条件
 合計料金(1部屋あたり消費税込み)
 下限 下限なし
 上限 上限なし
 検索

地図
 最近見た宿泊施設
 2軒の閲覧履歴があります(ページ 1/1)

施設紹介 プラン一覧 写真・動画(18) 地図・アクセス お客様の声(176) クーポン一覧

このクチコミ・お客様の声

総合評価
 ★★★★★ 3.39
 アンケート件数: 176件

評価内訳
 5点 17件
 4点 35件
 3点 22件
 2点 11件
 1点 6件

項目別の評価
 サービス ★★★★★ 3.06
 立地 ★★★★★ 3.28
 部屋 ★★★★★ 3.33
 設備・アメニティ ★★★★★ 3.00
 風呂 ★★★★★ 3.88
 食事 ★★★★★ 3.21

レビュー検索条件
 投稿の種類
 クチコミ (感想・情報) クチコミ (苦情)
 指定なし (147件)
 キーワード
 絞り込む
 投稿者
 一人 家族 恋人 友達 仕事仲間
 指定なし
 性別
 指定なし 男性 女性
 絞り込みを解除

最新投稿
 評価が高い順 (総合 | サービス | 立地 | 部屋 | 設備・アメニティ | 風呂 | 食事)
 147件中 1~20件表示 [1 | 2 | 3 | 4 | 5 | ... 全 8 ページ] 次の20件

投稿者さんのクチコミ (感想・情報)
 投稿者さん 2015年09月15日 16:25:21
 3歳の子どもと利用しました。室内に子ども用スリッパを用意してくれるなど、子どもにも優しい対応でした。ほっともっとフィールド神戸の野球観戦で利用しましたが、球場からも近く、ぜひまた利用したいです。
 レビューを評価してください 不適切なレビューを報告する
 このレビューは参考になりましたか? はい いいえ
 旅行の目的 ... レジャー
 同伴者 ... 家族
 宿泊年月 ... 2015年09月

2015年09月23日 13:13:20
 この度は、投稿者さんをご利用いただきまして、誠に有難うございました。
 近辺には球場以外にも、農業公園・総合運動公園・明石海峡大橋等様々な観光施設がございます。子供様にもお楽しみいただけるのではないかと幸いです。
 次回は是非とも観光にもご利用くださいませ。
 またのご利用をお待ちしております。

顧客レビュー文書
 販売者からの応答文書

ご利用の
 ご利用のお部屋 【『和洋室』ベッド・座敷でリフレッシュ】

クチコミを投稿する
 *クチコミを修正する

宿泊プラン一覧
 【イベントプラン】汗をかいたら温泉でさっぱりリフレッシュ☆宿泊まりプラン
 【観光運動公園プラン】汗をかいたら温泉でさっぱりリフレッシュ☆食事なしプラン
 【学会プラン】お勉強のあとに温泉でさっぱりリフレッシュ☆朝食なしプラン
 ☆当日限定☆朝食不要プラン
 【楽泊まり】プラン! 天然ラジウム温泉を満喫しよう☆
 【クオカード5000円付】+【きれいな天然温泉】でご利用プラン
 【クオカード1000円付】+【きれいな天然温泉】でご利用

図 1.2: EC サイト「楽天トラベル」におけるレビュー投稿と応答の例

書が不適切である場合は読み手の印象が悪化し (大沢他, 2010a)、対応に失敗してしまうことから、応答文書は重要な役割を持つ。特に、楽天トラベルなどが扱う無形のサービス財は有形の商材に比べると、サービス提供時の様々な諸条件の違いによって品質が大きく異なることからサービスの失敗が起りやすい (Lovell & Wright, 1999)、失敗への対応はより重要となる。

以上の背景を受け、本研究では商品紹介文、及びレビュー文書への応答文書を対象とした文書作成支援を目的とする。これらの文書はどのような内容が記載されているか、すなわち「商品紹介文において、顧客にとって魅力的な商品特長が被覆されているか」や「応答文書において、顧客からの質問・要望・苦情などに漏れなく回答が行われているか」が重要となる。そこで、これらの文書に対して何を書くべきかを提案することで、商品販売者への文書作成支援を実現する。具体的には、商品紹介文の場合は書き手に対して顧客にとって魅力的な商品特長を提示し、応答文書の場合は書き手に対して対応が漏れている文を提示することで、文書作成支援を行う。

このような「何を書くべきか」の提案による文書作成支援を実現するために、本研究ではレビュー文書を活用した手法を提案する。レビュー文書は顧客目線からの実体験に基づく情報を豊富に擁しており、文書作成支援にも有用な情報が多数含まれている。従来のレビュー文書を用いた研究は、レビュー情報の分析による情報の抽出・整理自体を目的とした研究が主であり (Liu, 2012; 乾・奥村, 2006)、レ

顧客 (レビュー文書作成者及びサイト閲覧者) の印象を向上させる点で有効であるとする立場を取る。

ビュー分析手法の応用は商品売上の予測 (Liu et al., 2007; Mishne & Glance, 2006) が中心であった。これに対し、本研究ではレビュー文書の分析手法を文書作成支援へ用いることで、レビュー文書の新たな応用可能性を示す。

本論文では、各商品の商品紹介文の作成支援を研究課題 1、応答文書の作成支援を研究課題 2 で扱う。以下、それぞれの研究課題について概略を示す。

研究課題 1: 商品紹介文に相応しい評価視点の提示 (第 3 章)

本研究課題では、商品検索画面・一覧画面などで提示される商品紹介文の作成支援を目的として、商品の紹介に相応しい商品特長を顧客レビュー分析に基いて提示する問題を提案し、解決手法を検討する。紹介文は顧客にとって最も魅力的な特長を記載することが求められるが、紹介文によっては顧客が求める特長が述べられていない場合や、紹介文で売り込む特長と顧客が求める特長が一致しない場合がある。このような販売者と顧客の間のギャップを埋めるため、本研究課題ではレビュー集合の活用を提案する。具体的には、対象商品と他商品集合の間で顧客レビュー集合を比較することで、商品紹介文として相応しい商品特長を抽出する。本研究課題を解決するため、本論文では評価視点表現と評価視点グループの対応付けと、評価視点グループのランキングにより行う枠組みを提案する。評価実験を通じ、提案した枠組みの有効性を示す。

研究課題 2: 対をなす二文書間における文対応関係の推定 (第 4 章)

本研究課題では、対をなす二文書間、例えば顧客からのメール文書やレビュー文書に対する応答文書における応答文の作成支援を目的として、対となる二つの文書間における文書を跨いだ文レベルでの対応関係を推定する問題を提案し、解決手法を検討する。対応関係の例として、質問と応答、苦情と謝罪、依頼と回答などがある。文対応関係の推定が実現すれば、応答文書の作成時に応答先文書の質問や苦情などへの対応漏れがないかを確認できる。特に、苦情への対応漏れは応答相手 (レビューの書き手) だけでなくサイト閲覧者の印象を悪化させることから (大沢他, 2010a)、対応漏れの検出が実現すればより確実な応答文書の作成が可能となる。本研究課題を解決するため、本論文では文種類に注目した推定手法を提案する。まず最初に、予め文種類を推定し、推定された文種類を文対応推定の素性として投入するモデルを示す。次に、文種類と文対応を同時に推定する統合モデルを示す。評価実験を通じ、提案した枠組みの有効性を示し、更に統合モデルでは文対応推定性能が向上したことを示す。

1.2 本論文の構成

本論文の以降の構成は次の通りである。まず、第 2 章では、文書作成支援に関連する研究および顧客レビュー文書を活用した研究について俯瞰する。第 3 章では商品紹介文を対象とし、商品紹介文として相応しい商品特長を抽出する課題と解決手法を提案する (研究課題 1)。第 4 章では、顧客からのレビューや問い合わせと応答文書の対を対象とし、文レベルでの対応を推定する課題と解決手法を提案する (研究課題 2)。最後に、第 5 章では本研究の貢献をまとめ、今後の課題について述べる。

第2章

関連研究

本章では、本研究と関連する先行研究を俯瞰する。まず、文書作成支援に関する研究について2.1節でまとめる。次に、顧客レビュー文書を利用する研究について2.2節でまとめる。最後に、これらの関連研究を受けた本研究の位置付けを2.3節で説明する。

2.1 文書作成支援に関する研究

文書作成支援に関する研究については、各分野において当該分野の研究を網羅した調査論文はいくつか存在するが、一方でより高次のレイヤーである「文書作成支援」という観点から研究を概括した文献は筆者の知る限りでは存在しない。そこで、本節では最初に文書作成支援研究の全体像を示し、続いて各分野の概要と動向について述べることで、文書作成支援研究についてまとめる。

2.1.1 文書作成支援の分類

文書作成支援研究の分類を文書作成の流れに従って示すために、最初に一般的な文書作成の過程を説明する。文書作成の過程と関連して、Flower & Hayes (1981) は文の作成過程が構想 (Plan)、言語化 (Translate)、検討 (Review) の繰り返しからなるとしている。これに対し、本研究では図 2.1 に示すように、文書全体と言う大局的な視点で作成過程を説明する。まず、文・文書の作成にあたり、どのような文や文書であれば目的が達成できるかを構想する (①文・文書の構想)。次に、構想を言葉の形で表現し、実際に文・文書を書く (②構想の文書化)。最後に、文書の構造・表現の見直しや誤字・脱字の修正を行う (③推敲・校正)。以上の過程を経て完成した文書は、最終的に発行・出版・公開されたり、特定の相手に送付されたり、または保存されたりすることで目的が果たされる。

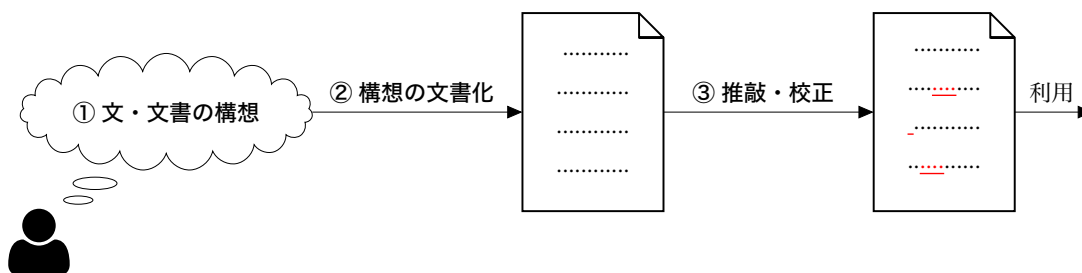


図 2.1: 文書作成の流れ

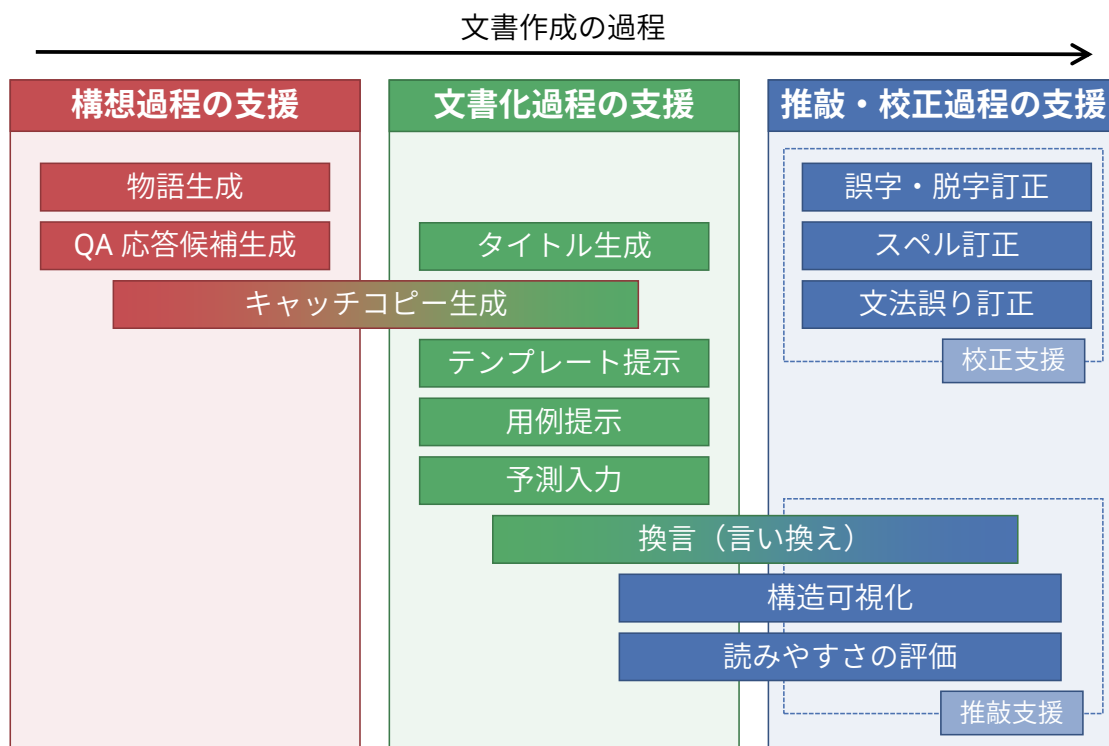


図 2.2: 文書作成支援の分類

これらの各過程ごとに、コンピュータにより様々な支援の試みが行われている。具体的に、文書作成支援を分類した図を図 2.2 に示す。構想過程の支援は、「何を書くべきか」の提案、すなわち目的を達成するための内容の提案による支援である。文書化過程の支援は、何を書くべきかが与えられている下で「どのように書くべきか」の提案、すなわち構想の言語化を行うために単語・文・文書やその断片などの提案を行う支援である。校正・推敲過程の支援は、完成した文書に対する誤字・脱字などの誤りや読みにくい文・文書などの検出、及び修正候補の提案による支援である。

以下、構想過程の支援、文書化過程の支援、校正・推敲過程の支援についてそれぞれ述べる。本研究は「何を書くべきか」を提案することによる文書作成支援が目的であるため、最も関連が深い構想過程の支援に関連した研究を重点的に説明する (2.1.2 節)。続いて、文書化過程の支援、校正・推敲過程の支援に関連した研究を概観する (2.1.3, 2.1.4 節)。

2.1.2 構想過程の支援

構想過程の支援は、何を書くべきかをユーザに提案することによる文書作成支援である。構想過程での支援は、目的とする文書の性質によって必要となる情報源や効果的な支援の態様が異なることから、目的文書によって異なる手法が用いられている。構想過程の支援を目的文書によって大別すると、以下の 2 つがある。

- (1) 物語・小説など、被読されること自体が目的となる文書 (物語生成)
- (2) 文書を通じた伝達・通信・コミュニケーションが目的となる文書 (Q&A 応答候補生成、キャッチコピー生成)

文書の性質により、(2)では(1)と異なり目的に応じた追加情報が用いられる。以下、(1)と関連する物語生成について、(2)と関連する Q&A 応答候補生成、キャッチコピー生成について、順次説明する。

物語生成

物語自動生成による物語制作支援の方式として、小方 (2007) は以下の 4 点を挙げている。

- (1) システムが粗筋や断片を作り出しユーザがそれを肉付けする方式
- (2) システムが多数の可能性を提示しユーザがそれを編集する方式
- (3) 多数生成された物語の系列を合成して多重構造の物語を作成する方式
- (4) 作品全体の骨組みは存在するがその部分を不確定なものとして残しそれをシステムが埋めて行く方法

明示的に物語制作支援を目的としている研究には佐久間・小方 (2005) があるが (方式 (1))、物語自動生成システムも方式 (2) として捉えれば制作支援と考えることができる。以下、物語自動生成システムに関する先行研究を述べる。

物語論によれば、物語の相は物語内容 (historie; 何を語るか)・物語言説 (récit; どのように語るか) として区別される。このうち、構想過程の支援と関連が深い物語内容の生成システムとしては、既存ストーリーの合成・再構築に基づく手法 (Callaway & Lester, 2001; 佐久間・小方, 2006) と、登場人物 (キャラクタ、エージェント) のシミュレーションに基づく手法 (Bringsjord & Ferrucci, 1999; Meehan, 1976; 鴨崎他, 2008; 吉崎他, 2010) がある。いずれの場合とも、物語は被読されること自体が目的であるため、構想に際して新たな事象を生成していると言う特徴を持つ。

QA 応答候補生成

質問応答 (Question Answering) は、自然言語での質問に対する回答を抽出・生成するシステムである。質問応答の出力はプログラム利用、例えば検索エンジンや対話エージェントなどが主であるが、人間がシステムの出力に基いて回答文書を作成するという用途においては構想過程の支援としても捉えることができる。以下、質問応答システムに関する先行研究を述べる。

質問応答で最も重要となるのは、回答に必要となる情報源である。初期には BASEBALL (Green Jr. et al., 1961) や LUNER (Woods, 1973) などのように、あるドメイン (BASEBALL は野球、LUNER は地質学) に特化した構造化データベースを用いる手法が提案されているが、これらの手法は構造的知識の構築コストが高いと言う問題があった。これに対し、近年では構造化されていない Web 文書から質問に関連した文書を探索し、回答を導く手法が提案されている (Katz et al., 2004, 2005; Radev et al., 2002)。例えば、Radev et al. (2002) は、複数のウェブ検索エンジンにより得られた文書から、質問と関連する文を抽出し、さらに回答となるフレーズを抽出する手法を提案している。

キャッチコピー生成

キャッチコピー生成は、商品と関連するテキスト集合を入力とする手法とキーワードを入力とする手法に大別できる。入力とするキーワードの要否を構想の要否と考えると、前者の手法は構想過程の支援と関わりが深いため、以下では前者の手法について扱う。後者については構想がキーワードとして予め与えられている下での文書化過程における支援とし、2.1.3 節で扱う。

商品の関連テキストを用いてキャッチコピーを生成する手法として、中野・鬼沢 (2008); 幅野・浦谷

(2013, 2014)の手法がある。中野・鬼沢(2008)は、商品説明文からキーワードを抽出し、共通した概念を持つ連想語句を用いたキャッチコピー生成手法を提案している。幅野・浦谷(2013, 2014)は、映画のあらすじとレビュー文書集合を用い、既存のキャッチコピーから抽出した文法構造を基に映画キャッチコピーを生成する手法を提案している。

2.1.3 文書化過程の支援

文書化過程の支援システムは、実際に文書を書く段階において構想の言語化を支援するためのシステムである。以下、文書化過程の支援を目的とした研究の概要と、主な研究についてそれぞれ述べる。

タイトル生成は、与えられた記事やウェブページなどのテキストから限られた長さのタイトルテキストを生成する技術である。まず、与えられた文書からタイトルとなる部分文字列を抽出する手法として、藤村他(2005); 南野・奥村(2005)の手法がある。藤村他(2005)はニュース記事を、南野・奥村(2005)はブログ記事を対象とし、HTMLの構造や本文中の単語の品詞・出現分布などに基いてタイトル文字列を抽出する手法を提案している。一方、記事中の文字列をそのまま用いるのではなく、動的に生成する手法として Banko et al.(2000)の手法がある。Banko et al.(2000)は、タイトルが付与されている記事集合を用い、記事の本文とタイトルに用いられる単語を対応付けるモデルと、タイトルの言語モデルを学習することにより動的な生成を実現する。2.1.2節で述べた構想支援におけるキャッチコピー生成システムは、タイトル生成を推し進め、キャッチコピーへ特化した手法と捉えることもできる。

文書化支援におけるキャッチコピー生成システムとして、ある入力キーワードに対するキャッチコピーを自動生成するシステムがある(Yamane & Hagiwara, 2015, 2013; 松平・萩原, 2005, 2004)。松平・萩原(2005, 2004)は、入力された単語の概念に基づき、遺伝的アルゴリズムにより多数のキャッチコピー候補を生成する手法を提案している。Yamane & Hagiwara(2015, 2013)は、既存のキャッチコピー集合から類似したテーマを持つキャッチコピーを選択し、構成する単語を入れ替えることによるキャッチコピー生成手法を提案している。

テンプレート提示システムは、文書の定型性に注目して再利用できるようにするシステムである。納富・内山(1991)はソフトウェアのマニュアルを対象とし、パラグラフを部品化・構造化して再利用する枠組みを提案している。また、恒川他(1993)は案内文や通知文などを対象とし、いくつかの変数(プレースホルダ)を持つ文書を予め用意し、文書作成時に各変数と対応付いた語句をユーザに選択させることで文書作成を容易にするシステムを提案している。

予測入力とは、打鍵^{*1}されたキー入力に基づき、一部の打鍵が省かれた状態からユーザが意図しているであろう入力内容を自動で推定するシステムである。日本語を対象とし、単語や文の入力完了を待つことなく未打鍵の部分を予測するシステムとして Nanashiki(小松他, 2003)、Social IME(奥野・萩原, 2009)、Mozc(工藤他, 2011)などがある。英語についても同様に Reactive Keyboardなどが提案されている(Darragh et al., 1990)。その他、携帯電話の multi-tap 方式^{*2}の連続打鍵を省略した single-tap 方式による入力単語の予測システムが提案されている(Smith & Goodwin, 1971)。

用例検索とは、実際に使用されている用例を検索するシステムである。特に、外国語文書の作成や論文・特許などについて、ユーザが当該ドメインにおける文書作成経験が少ない場合において、用例の提

^{*1} なお、厳密にはスマートフォンのフリック入力など物理キーの押下によらない入力方式も存在するが、ここでは文字を入力することを広義的に打鍵と呼ぶ。

^{*2} 携帯電話のように入力ボタンが限られている場合において、例えば“2”を1回打鍵することで“a”、続いて2回打鍵することで“b”、続いて3回打鍵することで“c”を出力するような入力方式。

示は文書化支援として効果がある。最も単純には、文書検索技術の向上とウェブ公開文書の増大が進んだ今日においては検索エンジンへのキーワード検索により用例を得られるが、これらの情報を更に整理・集約するシステムも提案されている (Tanaka-Ishii & Nakagawa, 2005)。また、キーワード入力ではなく、文を入力として受け取り、類似した構造を持つ文を提示するシステムも提案されている (Resnik et al., 2005; 兵藤他, 1996)。

換言 (言い換え) とは、単語や文の意味が同一となるような別表現を指す。換言の提案は文書作成中・作成後のいずれでも可能なため、文書化・推敲過程のいずれでも用いることができるが、本論文ではここで扱う。換言は機械翻訳や情報抽出の前処理などシステムの都合で行う場合もあるが、以下のように人間の都合で行う場合がある。例えば、読みやすさ・分かりやすさの評価 (高橋・牛島, 1991) に基づき、文をより読みやすい・分かりやすい文へ換言するシステムが提案されている (林・菊井, 1991)。また、表現の言い換えによる文・文書の簡素化 (Carroll et al., 1998)、字数削減や要約 (山崎, 1998; 若尾他, 1997) なども提案されている。

2.1.4 校正・推敲過程の支援

校正・推敲過程の支援システムは、既に完成している文書を対象として、校正や推敲を行うシステムである。なお、本研究では、校正と推敲を次のように区別して扱う。まず、校正とは文・文書の内容を変更せず、誤字・脱字などの修正を行う行為を指す。一方で、推敲とは表現や文・文書の構造などに踏み入って修正を行う行為を指す。

文書の校正・推敲支援による文書作成支援は、構想支援・文書化支援よりも比較的容易であり、加えてその需要の高さから古くから活発な研究が行われている分野である。以下、本節では校正支援、推敲支援の順に、研究の概要と主な研究を示す。

校正支援

校正支援は、英語などの西欧言語と日本語の言語特性の違いから異なる形で研究が進められてきたため、これらを分けて説明する。

まず、英語のスペル誤り訂正は、英文校正において最も重要な事項であることから、文書作成支援の中でも最初期に行われた試みである。英語の誤りは非単語誤り (non-word error; 実在しない単語への誤り) と実語誤り (real-word error; 実在する単語への誤り) に分けられる。最初は、非単語誤りに対し、Ralph E. Gorin が予め用意した辞書にない単語を誤りと指摘するプログラム SPELL を開発した (Peterson, 1980)。その後、辞書を用いる手法は実装が簡単である一方、適切な辞書サイズの設定が難しい点、実語誤りを検出できない点と言う 2 つの難点があるため、周辺文脈を用いる手法に感心が寄せられた。主な手法は、周辺文脈を用いた確率的モデルにより判定を行う手法である (Golding & Roth, 1999; Golding & Schabes, 1996; Jones & Martin, 1997; Mangu & Brill, 1997; Mays et al., 1991; Yarowsky, 1994)。正解候補となる単語の限定のために、どのように誤るか・誤りやすいかと言った Damerau (1964); Mitton (1996); Pollock & Zamora (1983) の知見も有用である。更にはスペル誤り訂正にとどまらず、文法誤りの検出や訂正を目的とした研究も行われている (Bustamante & León, 1996; Thurmair, 1990)。

これに対し、日本語は英語などと比較すると校正支援が難しい特徴を持つ。例えば、日本語が分かち書きされない言語であることから、英語のスペル誤りとは異なり、日本語の誤字・脱字の検出には辞書による手法が使えない。そのため、日本語校正支援は個別の現象へ特化し、明らかに誤りと判定可能な知識 (ルール) の利用による手法が提案された (下園他, 1994; 菅沼他, 1990; 菅沼・牛島, 1991; 鈴木・武

田, 1989; 脇田・奥村, 1993)。これらの手法は規則の網羅が難しい点や規則の衝突が問題となる点と言う難点があることから、その後、確率的に出現可能性が低い文字列を指摘する手法も提案されている(荒木他, 2002; 新納, 1999)。また、同音異義語(かな漢字変換)誤りのように誤りに対する正解候補が限定できる場合は、それらの候補との確率比に基づく手法もある(三品他, 2004)。

推敲支援

推敲支援としては、換言(2.1.3節)のように文を言い換える他、文書構造を可視化する手法や読みやすさを評価する手法がある。

文書構造の可視化によれば、文書の形では見えにくかった構造が明らかになるため、文書の推敲が容易になる。松本他(2009)は、論文などの技術文書を構成する文の修飾-非修飾関係を可視化する手法を提案した。また、石塚(2005)は英文記事を対象とし、名詞・固有名詞などの間の関連を可視化する手法を提案している。

読みやすさの評価は、ある読みやすさの指標に基いて文や文書にスコアを与える手法である。具体的に、英文書の読みやすさを評価する指標としては Dale-Chall Readability (Dale & Chall, 1948)、Flesch Reading Ease Score (Flesch, 1974)、Automated Readability Index (Senter & Smith, 1967)、Coleman-Liau index (Coleman & Liau, 1975)などが提案されている。これらは、文書中の語数や文の長さ、用いられている単語の語彙レベルなどに基いて評価を行っている。日本語文に対しては、文の長さに加えて建石他(1988)が字種(平仮名・片仮名・漢字・英字)の頻度などを用いた指標を提案している。また、石岡・亀田(2003)は小論文の自動採点システムにおいて読みやすさを評価する際、連体修飾の用言の数や連用形・接続助詞の区の並びの最大数なども用いた指標を用いている。

2.2 顧客レビュー文書を用いた研究

顧客レビュー文書を用いた研究は、主に評判分析 (Sentiment Analysis)・意見マイニング (Opinion Mining) と呼ばれる分野で行われている。評判分析、又は意見マイニングとは、商品 (product)・サービス (service) などに対する人々の意見 (opinion)・評判 (sentiment)・評価 (evaluation, appraisal)・態度 (attitude)・感情 (emotion) を分析する研究である (Liu, 2012)。これらの分野については Liu (2012); 乾・奥村 (2006) による調査が詳しい。本節では、Liu (2012); 乾・奥村 (2006) を参考に、顧客レビュー文書を用いた評判分析研究の流れを概観し、続いて評判分析手法を応用した研究を示す。

2.2.1 評判分析研究の概観

評判分析と関連して最も古くからある研究は、単語の極性 (Semantic Orientation; polarity) 推定である。ここで、単語の極性とは、肯定的・否定的の別を指す。Hatzivassiloglou & McKeown (1997) は単語極性の推定タスクを提案し、“and” や “but” などの接続詞で結ばれる語に注目した推定手法を提案した。また、Turney (2002) は相互情報量に基づく極性推定手法を提案している。

評価極性を軸に考えた場合、分類できる対象としては次の2つがある。一つは、評価が記述される文書である。このような文書の主な例としてはレビュー文書があるが、ブログ投稿やマイクロブログ投稿にも評価が記述されることがある。このような文書を、まとめて評価文書と呼ぶ。評価文書の分類について、Pang et al. (2002) は評価文書を肯定的・否定的の2カテゴリに分類する手法を示している。更に、Yu & Hatzivassiloglou (2003) はより細かい単位である文に対し、肯定的・否定的・極性なしの3カテ

リへの分類手法を提案している。

もう一つの分類対象は、評価の対象、すなわち商品やサービスである。商品やサービス全体の評価は、多くの場合はレビュー情報がレーティング（いわゆる5つ星評価など）を持つためそれらの集約により実現できる（なお、持たない場合は、先に上げた文書の極性判定と関連する問題である、文書のレーティング予測問題として捉えることができる。レーティング予測に対しては、Pang & Lee (2005) が SVM 回帰を用いた手法を提案している）。そのため、商品の部分や属性などのより詳細で、レーティングが与えられていない場合も多いものを対象とした分析へと関心が寄せられた。これら商品の部分や属性は、評判分析において評価視点 (aspect) と呼ばれ、評価視点を表す単語やフレーズは評価視点表現 (aspect expression) と呼ばれる (Liu, 2012)。評価視点に注目すると、ある商品ドメインにはどのような評価視点があるかを抽出・整理する課題、それぞれの評価視点の極性を推定する課題の2つがある。評価視点の極性判定は、評価視点表現を含む文に対し、先に挙げたレビュー文の分類手法を利用するなど対応できる。そこで、以下では評価視点の抽出・整理課題について詳しく説明する。

評価視点表現を抽出する手法としては、Hu & Liu (2004); Popescu & Etzioni (2005) による高頻度の名詞・名詞句に注目した手法が提案されている。しかし、これら評価視点表現を列挙するのみでは多くの表現が得られてしまうことから、評価視点表現のグループ化・ランキングによって整理する手法が提案された。所与の評価視点グループへ評価視点表現をグループ化する課題は Carenini et al. (2005) により提案され、併せて WordNet の同義語情報などに基づいたグループ化手法を提案している。また、Zhai et al. (2010) は、半教師あり学習に基づいたグループ化手法を示している。また、評価視点グループを与えずにレビュー文書集合から自動で構築する手法が Titov & McDonald (2008a) により提案されている。一方で、評価視点表現を与えられたドメインにおける重要度に応じてランキングすることで、評価視点表現を整理する課題が Zhang et al. (2010) により提案され、Web ページのランキング手法である HITS アルゴリズム (Kleinberg, 1999) の応用による手法を示している。これに対し、ドメイン全体ではなく、与えられた商品における特徴度合いに基いてランキングする課題が乾他 (2013) により提案され、対数尤度比 (Dunning, 1993) に基づく手法を併せて示している。

2.2.2 評判分析の応用

評判分析はレビュー文書に含まれる意見の整理・集約が目的であり、これらの情報はそれ自体で顧客や販売者にとって有用な情報であるが、これらの情報を更に別の課題に応用する研究もある。本節では、このような評判分析の応用研究について述べる。

直接的な応用としては、商品の評判情報と密接な関わりを持つ商品売上の予測がある。Liu et al. (2007); Mishne & Glance (2006) はブログ文書を用い、いずれも評判分析を利用して映画の売上予測を行っている。また、近年ではマイクロブログ文書を用いた研究も行われている。Asur & Huberman (2010) は、Twitter^{*3}投稿への評判分析の適用による映画の売上予測手法を提案している。これらの研究を更に推し進めたものとしては、株価予測 (Oh & Sheng, 2011) や選挙結果の予測 (Tumasjan et al., 2010) がある。

その他の応用としては、夢テキストに対する評価情報の自動付与 (Razavi et al., 2014)、メール文書における性差の分析 (Mohammad & Yang, 2011)、チャット上での性犯罪者の検出 (Bogdanova et al., 2012)、うつ傾向を持つマイクロブログユーザの推定 (Tsugawa et al., 2015) などがある。

*3 Twitter: <https://twitter.com/>

2.3 本研究の位置付け

本論文で提案する課題は、いずれも「何を書くべきか」の提案を、顧客レビュー文書を利用することで実現する課題である。そこで、最初に構想過程の支援を目的とした研究の中での位置付けを説明し、続いてレビュー文書を利用する研究の中での位置付けを説明する。

本研究の文書作成支援の対象は EC サイトに関連したテキスト、具体的には商品紹介文・販売者応答文である。これらの文書を対象とした構想過程の支援に関する研究は、2.1.2 節で挙げたキャッチコピー生成に関連があるものの、その他については筆者の知る限り行われていない。そのため、本研究は構想過程における作成支援が可能となる、新たな文書領域を提案するものである。

レビュー文書を用いた従来の研究は、レビュー文書に含まれる意見の整理・集約自体が主な目的であった。2.2.2 節では、整理・集約された情報を応用する研究をいくつか挙げたが、本論文では第3章および第4章で詳細を述べるように、新たに文書作成支援への利用を提案する。これにより、レビュー文書の新たな応用可能性を示し、ユーザ生成コンテンツの利用価値を更に高めることができる。また、要素技術に注目すると、研究課題1ではレビューを構成する文へのラベル付与（文種類分類）、研究課題2では評価視点表現の構造化・ランキング手法を用いており、いずれも従来積み重ねられてきた評判分析手法に依拠している。これらの分析手法を用いた応用研究を示すことで、評判分析研究の価値を更に高めることに繋がると考えている。

第3章

商品紹介文に相応しい評価視点の提示

本研究課題では、商品検索画面や一覧画面などで提示される商品紹介文の構想過程における支援を目的として、商品の紹介に相応しい商品特長を顧客レビュー分析に基づいて提示する課題の提案と解決手法の検討を行う。本章では、最初に研究課題の背景と概要について説明する。次に、商品特長を評価視点及び評価視点グループの概念により定義づける。続いて、提案手法の説明に必要な基盤技術を導入する。次いで、本研究課題を解くための枠組みと、導入した基盤技術を利用した手法の提案を行う。更に、評価実験を通じて、提案した枠組みの有効性を示す。最後に、本研究の目的と関連した研究の概要と、本研究との違いについて述べる。

3.1 研究課題の背景と概要

EC は新たな販売チャネルを切り拓いただけでなく、顧客が商品の購買に至るプロセスにも大きな変化を与えた。EC 登場以前の購買モデルとして、1898 年に E. St. Elmo Lewis は“attract attention, maintain interest, create desire”モデルを提案し、後に“get action”を追加した AIDA モデルを提案している (Strong, 1925)。Lewis によれば、消費者の購買行動プロセスは、商品の存在を認知し (Attention; 注目)、関心が生まれ (Interest; 興味・関心)、購買欲求が高まることで (Desire; 欲望)、購入へと至る (Action; 購買) と言う 4 段階で説明される。これに対し、EC の登場と普及による購買行動プロセスの変化を踏まえ、新たに AISAS、AISCEAS、AISDALSLove などのモデルが提案されている。AISAS モデルは株式会社電通が提案したモデルであり (現代用語の基礎知識, 2015)、Attention (注目)、Interest (興味・関心) の後に Search (検索) が、Action (購買) の後に Share (共有) が起こるとしている。また、AISCEAS モデルは有限会社アンヴィコミュニケーションが提案したモデルであり (アンヴィコミュニケーションズ, 2015)、Attention (注目)、Interest (興味・関心) の後に Search (検索)、Comparison (比較)、Examination (検討) の各段階があり、Action (購買) の後に Share (共有) があると考えられる。更に AISDALSLove は Wijaya (2012) が提案したモデルであり、Attention (注目)、Interest (興味・関心) に続いて Search (検索) が、Action (購買) の後に Like/dislike (好悪判断)、Share (共有)、Love/hate (忠誠・嫌悪) 段階があるとしている。これら全てのモデルで Action (購買) の前にウェブ上での Search (検索) 行動が起こると考えており、商品販売者にとって検索段階の広告戦略は重要と言える。

検索段階において商品のプレゼンスを高めるために重要となるのは、商品紹介文である。商品紹介文は検索クエリ語の一致に用いられると同時に、消費者が検索結果を更に選別する際にも用いられており、システムとユーザの両者から利用される。ここで、商品紹介文に重要となるのは以下の 2 点である。

- (1) 訴求する特長が顧客にとって重要視されているか
- (2) 訴求する特長が他商品と差別化できているか

従来、(1), (2) を満足させるような商品特長の選定は人手で行われていた。このような特長の選定にあたっては、レビュー集合には顧客の視点で見た際の重要な特長に関する言及が含まれているため、他商品を含めた顧客レビュー集合が有用である。しかし、他商品を含めたレビュー集合の人手による分析は、レビュー集合の規模が大きい場合は必ずしも容易ではない。そのため、レビュー集合の分析に基づいた(1), (2) を満たす特長の自動抽出は、商品紹介文の書き手にとって有用である。

以上の背景を踏まえ、本研究課題では商品紹介文の構想過程における文書作成支援を目的として、商品の紹介に相応しい商品特長を提示する課題を提案する。先に商品紹介文に重要な2要素を挙げたが、本研究ではこれらを満たす特長を、顧客レビュー集合を情報源として利用することで順位付けする。レビュー集合を用いれば、要素(1)を満たす特長が抽出できる。また、他商品の全レビュー集合との比較を通じ、要素(1), (2)の双方を満たす特長に限定できる。これらの抽出された特長を商品紹介文の書き手に提示することで、商品紹介文の作成支援を実現する。

3.2 評価視点・評価視点グループ

3.2.1 評価視点

以上の説明では商品の属性などを指す際に「特長」という語を用いてきたが、本研究課題では**評価視点** (aspect) 及び**評価視点表現** (aspect expression) という概念を用い、より厳密に定義する。これらの定義には、以下に示す Zhang & Liu (2014) による “aspect and aspect expression” の定義を用いる。なお、以下の定義におけるエンティティとは、商品やサービスなどを指す。

定義 1 (評価視点・評価視点表現 (Zhang & Liu, 2014)). エンティティ e の**評価視点** (aspect) とは、 e の部分や属性を指す。**評価視点表現** (aspect expression) とは、評価視点を指し示す実際の単語やフレーズを指す。

具体的な評価視点の例として、Zhang & Liu (2014) は商品が携帯電話の場合、商品の部分として「バッテリー」「画面」、商品の属性として「音質」「大きさ」「重さ」などがあるとしている。また、評価視点「音質」に対応する評価視点表現として「音 (sound)」「音声 (voice)」や「音質 (voice quality)」自身を挙げている。本研究の実験の場合、データとして楽天トラベル (商品は宿泊サービス) を用いているため、具体的な評価視点の例としては「立地」「景観」「食事」や「従業員の態度」「風呂のシャワー」などがある。

3.2.2 評価視点グループ

続いて、本研究課題では評価視点をグループ化した**評価視点グループ**を定義する。

定義 2 (評価視点グループ). **評価視点グループ** (aspect group) とは、評価視点からなる集合を指す。**評価視点グループを構成する表現**とは、評価視点グループを構成する各評価視点と対応する評価視点表現集合の和集合を指す。

評価視点グループの構成方法は対象とする商品や目的によって異なるが、本研究では意味的なまとま

りによって評価視点グループを構成する。本研究の評価実験に用いる具体的な評価視点グループの一覧は3.4.3節で示すが、宿泊サービスの場合は「立地」「サービス」「風呂」などがある。評価視点グループ「サービス」は評価視点「フロントの対応」や「ポーターの対応」などを内包し、「風呂」は「風呂のシャワー」や「風呂の温度」などを内包する。さらに、評価視点グループ「サービス」は評価視点「フロントの対応」を内包し、評価視点「フロントの対応」は評価視点表現「チェックイン対応」「チェックアウト対応」などを内包することから、評価視点グループ「サービス」は「チェックイン対応」「チェックアウト対応」を評価視点グループを構成する表現として持つ。

3.3 基盤技術

本研究課題に対して提案する枠組みは、評価視点表現と評価視点グループの対応付けと、評価視点グループのランキングから構成される。評価視点表現と評価視点グループの対応付けは Zhai et al. (2010) の提案した評価視点表現と評価視点の対応付け手法に、評価視点グループのランキングは対数尤度比 (Log-Likelihood Ratio; LLR) (Neyman & Pearson, 1928) に基づいて行う。そこで、本節では Zhai et al. (2010) の提案手法と対数尤度比について順次説明する。

3.3.1 評価視点表現の評価視点グループへの対応付け

評価視点表現と評価視点の対応付けを行う先行研究に Zhai et al. (2010) の手法がある。彼らは Nigam et al. (2000) による単純ベイズアルゴリズムを拡張した半教師あり文書分類手法を応用することで、小規模の評価視点表現と評価視点の対応付けシードと大量のレビュー文書集合から評価視点表現と評価視点の対応付けを実現している。

以下、Nigam et al. (2000) による半教師あり文書分類手法を説明し、続いて Zhai et al. (2010) による評価視点表現と評価視点の対応付け手法を説明する。

半教師あり文書分類 (Nigam et al., 2000)

文書分類問題とは、文書 d を入力とし、予め定義されたラベル集合 $C = c_1, \dots, c_{|C|}$ から最も適切なラベル $y \in C$ を出力する問題である。まず、Nigam et al. (2000) による半教師あり文書分類手法を示す前に、彼らの手法の基となる教師あり文書分類手法である単純ベイズ法 (Naïve Bayes) から説明する。

単純ベイズ法では、文書中の各単語が独立に生起していると仮定する。この仮定を利用すると、文書 d がラベル $y \in C$ である同時確率 $P(d, y)$ は以下の通り表せる。

$$P(d, y) = P(y)P(d|y) \quad (3.1)$$

$$\simeq P(y) \prod_{d_k} P(w_{d_k}|y) \quad (3.2)$$

ここで、 w_{d_k} は文書 d における k 番目の単語を表す。このモデルにおけるパラメータは $P(y)$ と $P(w|y)$ の2種類であり、それぞれ

$$\theta(y) \equiv P(y), \theta(w, y) \equiv P(w|y) \quad (3.3)$$

とおく。以下、便宜的にこれらのパラメータをまとめて θ とする。

以下、最初に単純ベイズ法のパラメータ学習方法を説明し、その後に学習結果に基づいた分類方法について説明する。単純ベイズ法では、学習データにおける文書とラベルの組がそれぞれ $\langle d^{(i)}, y^{(i)} \rangle$ で与

えられた際、学習データの同時確率を最大化するように θ を設定する最尤推定を用いる。最尤推定によれば、尤もらしいパラメータ $\hat{\theta}$ は以下のようにして求められる。

$$\hat{\theta} = \operatorname{argmax}_{\theta} \prod_i P(d^{(i)}, y^{(i)}; \theta) \quad (3.4)$$

式 (3.4) の $\prod_i P(d^{(i)}, y^{(i)})$ を対数化した対数尤度関数 $L(\theta)$ を求め、式 (3.2) を代入すると

$$\begin{aligned} L(\theta) &= \prod_i P(d^{(i)}, y^{(i)} | \theta) \\ &= \sum_i \left\{ \log \theta(y^{(i)}) + \sum_{d_k} \log \theta(w_{d_k}, y^{(i)}) \right\} \end{aligned} \quad (3.5)$$

である。ここで、学習データ全体を見た時にラベル c を持つデータの数を $N(c)$ 、ラベル c を持つ全文書での単語 w の出現回数を $N(w, c)$ とおくことにすると、式 (3.5) は次のようにできる。

$$L(\theta) = \left\{ \sum_c N(c) \log \theta(c) + \sum_c \sum_{w \in \mathcal{V}} N(w, c) \log \theta(w, c) \right\} \quad (3.6)$$

なお、 \mathcal{V} は全文書における単語の集合を示す。この $L(\theta)$ を最大化する θ を求めることで、最尤推定による θ を求められる。ここで、 $\theta(y)$ 、 $\theta(w, y)$ はそれぞれ確率パラメータであることから以下が成り立つ。

$$\sum_c \theta(c) = 1 \quad (3.7)$$

$$\sum_{w \in \mathcal{V}} \theta(w, c) = 1 \quad \forall c \in C \quad (3.8)$$

まず、 $\theta(c)$ と $\theta(w, c)$ は独立であることから $\theta(c)$ に注目すると、式 (3.6) で関連する項は

$$\sum_c N(c) \log \theta(c) \quad (3.9)$$

である。式 (3.9) を制約式 (3.7) の元で最大化する問題は、ラグランジュの未定乗数法によって解析的に解くことが可能である。ラグランジュ乗数を λ とおくと、式 (3.9) に対応するラグランジュ関数 $\mathcal{L}_{\theta(c)}$ は

$$\mathcal{L}_{\theta(c)} = \sum_c N(c) \log \theta(c) - \lambda \left(\sum_c \theta(c) - 1 \right) \quad (3.10)$$

とできる。今、ある特定の $c' \in C$ に注目する場合、 $\mathcal{L}_{\theta(c)}$ が最大となるための $\theta(c')$ は以下の式を解くことで求められる。

$$\frac{\partial}{\partial \theta(c')} \mathcal{L}_{\theta(c)} = \frac{N(c')}{\theta(c')} - \lambda = 0 \quad (3.11)$$

$\theta(c')$ について整理すると

$$\theta(c') = \frac{N(c')}{\lambda} \quad (3.12)$$

である。更に式 (3.7) を利用すれば $\lambda = \sum_c N(c)$ であることから、最終的にパラメータ $\theta(c')$ は以下の式により求めることができる。

$$\theta(c') = \frac{N(c')}{\sum_c N(c)} \quad (3.13)$$

同様に、ある特定の $c' \in C, w' \in V$ に対応する $\theta(w', c')$ に注目する場合、式 (3.6) の項

$\sum_c \sum_{w \in V} N(w, c) \log \theta(w, c)$ を制約 (3.8) の元で最大化する $\theta(w', c')$ は以下の式により求めることができる。

$$\theta(w', c') = \frac{N(w', c')}{\sum_{w \in V} N(w, c')} \quad (3.14)$$

続いて、単純ベイズ法において新たな文書が与えられた際の分類方法について説明する。新たな文書 d に対するラベルの分布は $P(y|d)$ であることから、 y があるラベル c となる確率は $P(y = c|d)$ である。推定したパラメータ $\hat{\theta}$ を利用した $P(y = c|d; \hat{\theta})$ に対してベイズの定理を適用すると、

$$\begin{aligned} P(y = c|d; \hat{\theta}) &= \frac{P(c; \hat{\theta})P(d|c; \hat{\theta})}{P(d; \hat{\theta})} \\ &= \frac{P(c; \hat{\theta}) \prod_{d_k} P(w_{d_k}|c; \hat{\theta})}{Z_d} \end{aligned} \quad (3.15)$$

とできる。ここで Z_d は正規化項であり $Z_d = \sum_{c_r} P(c_r; \hat{\theta}) \prod_{d_k} P(w_{d_k}|c_r; \hat{\theta})$ である。 $P(c; \hat{\theta})$ には式 (3.13) で求めた $\theta(c)$ を、 $P(w_{d_k}|c; \hat{\theta})$ には式 (3.13) で求めた $\theta(w, c)$ を用いれば良い。分類の際に d に対するラベルを一意に定める必要がある場合は、事後確率が最も高い $\operatorname{argmax}_c P(c|d; \hat{\theta})$ を選択すれば良い。

単純ベイズ分類器の導入が完了したので、ここから Nigam et al. (2000) による半教師あり文書分類手法について説明する。半教師ありの場合、ラベルが既知の文書と未知の文書が与えられるため、ラベルが既知の文書集合を \mathcal{D}^l 、未知の文書集合を \mathcal{D}^u として区別し、全文書集合を \mathcal{D} とする。また、ラベルが既知の文書 $d_i \in \mathcal{D}^l$ のラベルを $y_i \in C$ とする。ラベルが既知の文書集合の場合、文書集合全体の生起確率は式 (3.1) と同様である。一方、ラベルが未知の文書集合の場合は

$$P(\mathcal{D}^u) = \prod_{d_i \in \mathcal{D}^u} \sum_{c_j} P(c_j)P(d_i|c_j) \quad (3.16)$$

である。式 (3.1), (3.16) より、パラメータ θ の下での文書集合全体の生起確率 $P(\mathcal{D}; \theta)$ は以下のようになる。

$$\begin{aligned} P(\mathcal{D}; \theta) &= P(\mathcal{D}^u; \theta) \cdot P(\mathcal{D}^l; \theta) \\ &= \left\{ \prod_{d_i \in \mathcal{D}^u} \sum_{c_j} P(c_j; \theta)P(d_i|c_j; \theta) \right\} \cdot \left\{ \prod_{d_i \in \mathcal{D}^l} P(y_i)P(d_i|y_i) \right\} \end{aligned} \quad (3.17)$$

確率の最大化を考えるため、対数化すれば

$$\log P(\mathcal{D}; \theta) = \left\{ \sum_{d_i \in \mathcal{D}^u} \log \sum_{c_j} P(c_j; \theta)P(d_i|c_j; \theta) \right\} + \left\{ \sum_{d_i \in \mathcal{D}^l} \log (P(y_i)P(d_i|y_i)) \right\} \quad (3.18)$$

である。式 (3.18) には対数和を含む項があるため解析解を得ることは困難であるが、このような問題に対しては潜在変数を仮定し、Expectation-Maximization (EM) アルゴリズム (Dempster et al., 1977) を用いることで局所解を求めることができる。今、一般に観測変数が \mathbf{x} 、潜在変数が \mathbf{z} であり、同時分布 $P(\mathbf{x}, \theta)$ が与えられている下での尤度関数 $P(\mathbf{x}|\theta)$ を θ について最大化するための EM アルゴリズムの手順は以下の通りである。

1. パラメータの初期値を θ^{old} とする。

2. (E-ステップ) $P(z|\mathbf{x}, \theta^{\text{old}})$ を計算する。
3. (M-ステップ) $\theta^{\text{new}} = \operatorname{argmax}_{\theta} \sum_z P(z|\mathbf{x}, \theta^{\text{old}}) \log P(\mathbf{x}, z|\theta)$ を計算する。
4. 収束判定を行い、収束条件を満たしていなければ $\theta^{\text{old}} \leftarrow \theta^{\text{new}}$ とし、ステップ2に戻る。

そこで、EM アルゴリズムを適用するために、次のような潜在変数 z を仮定する。

$$z_{ij} = \begin{cases} 1 & \text{if } y_i = c_j \\ 0 & \text{otherwise} \end{cases} \quad (3.19)$$

潜在変数 z を用いれば、式 (3.18) は以下のように書き直せる。

$$\log P(\mathcal{D}, z; \theta) = \sum_{d_i \in \mathcal{D}} \sum_{c_j} z_{ij} \log (P(c_j|\theta)P(d_i|c_j; \theta)) \quad (3.20)$$

(対数) 同時分布が式 (3.20) で与えられている際、EM アルゴリズムを適用した場合の E-ステップと M-ステップの計算方法は以下になる。ここで、観測変数は \mathcal{D} 、潜在変数は z であり、パラメータは θ である。まず、E-ステップの計算は $P(z|\mathcal{D}, \theta^{\text{old}})$ はパラメータ θ^{old} のもとで文書集合 \mathcal{D} をそれぞれ分類する手順に相当する。これはパラメータを $\hat{\theta} = \theta^{\text{old}}$ とした際の式 (3.15) の計算に相当する。また、M-ステップの計算は現在の分類結果から θ^{new} を新たに推定する手順に相当する。これは式 (3.13), (3.14) の計算に相当する。

以上の手順をまとめると、半教師あり文書分類は、単純ベイズ法における学習結果の利用 (3.15) と学習 (3.13), (3.14) の繰り返しにより実現できる。なお、初期値について、Nigam et al. (2000) はラベルが既知のデータ \mathcal{D}^l のみから学習した値を与えている。本研究でも同様の初期値を用いる。

評価視点表現の評価視点への対応付け (Zhai et al., 2010)

Zhai et al. (2010) の提案した評価視点表現の分類は、各評価視点表現に対する擬似文書を作成し、Nigam et al. (2000) の手法を適用することで実現する。擬似文書は、注目している評価視点表現について、全文書集合における各出現時の前後数単語を集約することで作成する。また、擬似文書の分類において、文書ラベルは対応する評価視点を表す。

擬似文書の作成方法について、具体的に例を挙げて説明する。例えば、評価視点表現「風呂」に注目した際、全文書集合において「風呂」を含む文が以下の2つであった場合を考える (説明の都合上、単語に分割している)。

- 足を伸ばせる 風呂 に入れて良かったです。
- 大きなお 風呂 に入れてゆったり過ごせました。

今、文脈として考える単語数を Zhai et al. (2010) と同様に前後3単語とした場合、stop-word*¹ を無視した前後3単語は次の通りである。

- 足, 伸ばせ, 風呂, 入れ, 良かつ
- 大きな, 風呂, 入れ, ゆったり, 過ごせ

これらの単語から、bag-of-words により擬似文書を構成する。実際に評価視点表現「風呂」と対応する擬似文書 $d_{\text{風呂}}$ は以下の通りである。

$$d_{\text{風呂}} = \langle \text{足, 伸ばせ, 風呂, 入れ, 良かつ, 大きな, 風呂, 入れ, ゆったり, 過ごせ} \rangle$$

*¹ 本研究では、品詞が「助詞」「助動詞」「記号」の単語を stop-word とした。

以上のようにして擬似文書を全ての評価視点表現に対して作成する。その後、対応が分かっている評価視点辞書シード中の評価視点表現と対応する擬似文書をラベル付き、それ以外の評価視点表現と対応する擬似文書をラベルなしと考え、Nigam et al. (2000) の手法を適用する。結果として得られた各擬似文書のラベル、すなわち評価視点对応する評価視点表現へ付与することで、評価視点表現の評価視点への対応付けが完了する。

3.3.2 対数尤度比

対数尤度比 (Log-Likelihood Ratio; LLR) とは、Neyman & Pearson (1928) が提案した統計的指標である。対数尤度比は先行研究において、ある二つの単語間の関係度合いや、ある単語 w が与えられた二群のテキスト集合のいずれかにおける特徴度合いの指標として用いられている (Dunning, 1993)。具体例としては、単語のコロケーション抽出 (Manning & Schütze, 1999) や、分野特徴単語の抽出 (内山他, 2004)、古語テキスト間での特徴語抽出 (小林・小木曾, 2013)、特徴的評価視点表現の抽出 (乾他, 2013) などの応用例がある。

以下、対数尤度比の導出方法と二群のテキスト集合における特徴度合いの指標として適用する方法を、内山他 (2004) の解説を参考に説明する。今、二群のテキスト集合を A, B とし、ある単語 w に注目する場合を考える。ここで、二群のテキスト集合における単語を順次見た際、観測された単語を v について、次の二つの確率変数を導入する。

$$X = \begin{cases} 1 & \text{if } v = w \\ 0 & \text{otherwise} \end{cases} \quad (3.21)$$

$$Y = \begin{cases} 1 & \text{if 単語 } v \text{ がテキスト集合 } A \text{ で出現した} \\ 0 & \text{if 単語 } v \text{ がテキスト集合 } B \text{ で出現した} \end{cases} \quad (3.22)$$

これらの変数を用いれば、データ全体を

$$\mathcal{D} = \{ \langle x_1, y_1 \rangle, \dots, \langle x_{|\mathcal{D}|}, y_{|\mathcal{D}|} \rangle \} \quad (3.23)$$

と要約できる。ある仮説 H が与えられた下での \mathcal{D} の生起確率は

$$P(\mathcal{D}|H) = \prod_i P(X = x_i, Y = y_i|H) \quad (3.24)$$

で与えられる。今、仮説として以下の2つを仮定する。

H_{indep} : 確率変数 X, Y は互いに独立である

H_{dep} : 確率変数 X, Y は互いに従属である

対数尤度比は、これらの仮設の下でのデータ全体 \mathcal{D} の生起確率として与えられる。すなわち、LLR は以下ようになる。

$$\text{LLR} = \log \frac{P(\mathcal{D}|H_{\text{dep}})}{P(\mathcal{D}|H_{\text{indep}})} \quad (3.25)$$

式 (3.24) を代入すれば

$$\begin{aligned} \text{LLR} &= \log \frac{\prod_i P(X = x_i, Y = y_i|H_{\text{dep}})}{\prod_i P(X = x_i, Y = y_i|H_{\text{indep}})} \\ &= \sum_i \log \frac{P(X = x_i, Y = y_i|H_{\text{dep}})}{P(X = x_i, Y = y_i|H_{\text{indep}})} \end{aligned} \quad (3.26)$$

である。ここで、確率変数 X, Y はそれぞれ $X \in \{0, 1\}, Y \in \{0, 1\}$ の二値であることから、 X, Y をそれぞれ陽に列挙すれば展開できる。ここで、以下のような変数 a, b, c, d を導入する。

$$\begin{aligned} a &= \{X = 1, Y = 1\} \text{ の事象数} \\ b &= \{X = 1, Y = 0\} \text{ の事象数} \\ c &= \{X = 0, Y = 1\} \text{ の事象数} \\ d &= \{X = 0, Y = 0\} \text{ の事象数} \\ n &= \text{全事象数 } (a + b + c + d) \end{aligned}$$

これらの変数を用いれば、式 (3.26) は次のようにできる。

$$\begin{aligned} \text{LLR} &= a \log \frac{P(X = 1, Y = 1|H_{\text{dep}})}{P(X = 1, Y = 1|H_{\text{indep}})} + b \log \frac{P(X = 1, Y = 0|H_{\text{dep}})}{P(X = 1, Y = 0|H_{\text{indep}})} \\ &\quad + c \log \frac{P(X = 0, Y = 1|H_{\text{dep}})}{P(X = 0, Y = 1|H_{\text{indep}})} + d \log \frac{P(X = 0, Y = 0|H_{\text{dep}})}{P(X = 0, Y = 0|H_{\text{indep}})} \end{aligned} \quad (3.27)$$

さらに、 H_{indep} の元では X, Y は独立であるから、 $x \in \{0, 1\}, y \in \{0, 1\}$ とすると

$$P(X = x, Y = y|H_{\text{indep}}) = P(X = x|H_{\text{indep}}) \cdot P(Y = y|H_{\text{indep}}) \quad (3.28)$$

が成り立つ。式 (3.28) を式 (3.27) に代入した際に用いられている各確率変数の値を最尤推定すると、それぞれ以下の通りである。

$$\begin{aligned} P(X = 1|H_{\text{indep}}) &= \frac{a + b}{n} & P(X = 0|H_{\text{indep}}) &= \frac{c + d}{n} \\ P(Y = 1|H_{\text{indep}}) &= \frac{a + c}{n} & P(Y = 0|H_{\text{indep}}) &= \frac{b + d}{n} \\ P(X = 1, Y = 1|H_{\text{dep}}) &= \frac{a}{n} & P(X = 0, Y = 1|H_{\text{dep}}) &= \frac{b}{n} \\ P(X = 1, Y = 0|H_{\text{dep}}) &= \frac{c}{n} & P(X = 0, Y = 0|H_{\text{dep}}) &= \frac{d}{n} \end{aligned}$$

これらの値を用いれば式 (3.26) は

$$\begin{aligned} \text{LLR} &= a \log \frac{an}{(a + b)(a + c)} + b \log \frac{bn}{(a + b)(b + d)} \\ &\quad + c \log \frac{cn}{(c + d)(a + c)} + d \log \frac{dn}{(c + d)(b + d)} \end{aligned} \quad (3.29)$$

とできる。

但し、式 (3.29) に示す式では、与えられた単語 w が、テキスト集合 A, B のどちらで特徴的なかを区別できない。そこで、式 (3.29) に次の補正を加えることで、 A, B のどちらで特徴的なか区別できるようにする (影浦, 1997)。

$$\text{LLR}' = \begin{cases} +\text{LLR} & \text{if } (ad - bc) > 0 \\ -\text{LLR} & \text{otherwise} \end{cases} \quad (3.30)$$

この補正により、 LLR' は単語 w がテキスト集合 A で特徴的なほど正の高い値を取り、テキスト集合 B で特徴的なほど負の低い値を取るようになる。

対数尤度比を二群のテキスト集合における特徴度合いの指標として用いる利点として、以下の二点がある。まず、対数尤度比は単語 w の特徴度合いをテキスト集合 A の頻度だけでなく B の情報も用いて

いるので、 B との比較を通じて A において特徴的かどうかを測れる点がある。これにより、単語 w が A において他の単語と比較して高頻度で出現している場合、 B でも同様に高頻度で出現している場合は低いスコアが与えられる。加えて、類似の尺度である自己相互情報量では w がテキスト集合 A, B いずれにおいても低頻度な場合に特徴度合いを過大評価する欠点があるが (Manning & Schütze, 1999; 内山他, 2004)、対数尤度比は頻度情報もある程度考慮したスコア付けができる利点もある。

3.4 提案手法

本節では、3.2 節で説明した評価視点及び評価視点グループの概念を利用し、提案手法について説明する。まず、本研究課題の定式化を行い、課題を解決するための枠組みを提案する。続いて、本枠組みを構成する手法である、評価視点表現の評価視点グループへの対応付け・評価視点グループのランキングを行う各手法について順次説明する。

3.4.1 提案手法の枠組み

本論文では、本研究課題を「ある商品 h を入力とし、評価視点グループのリストを商品 h の紹介文として相応しい順に出力する」問題と定義する。商品販売者は、提示された評価視点グループのリストに基いて商品紹介文を作成することで、商品紹介文をより容易に作成できるようになる。本研究では、相応しい評価視点グループを判定する情報源として、レビュー文書集合を用いる。

ここで、提示する内容として評価視点表現や表現視点ではなく評価視点グループを用いる理由は、意味的に類似している評価視点表現や評価視点を同一に扱うためである。具体的に、商品が宿泊サービスの場合において、異なるレビュー文書で以下のように述べられている場合を例に説明する。

- 立地は抜群だと思います。
- 駅からの距離は徒歩1分と、とても便利でした。
- 雨でも傘なしで行ける道のりで助かりました。

これらのレビュー文は、異なる評価視点表現を用いており、また評価視点の定義に照らせば異なる属性を示している。しかし、いずれの表現でも、宿泊施設の立地について述べている点で共通している。これらを別個に扱おうと、類似した評価視点表現や評価視点が同時かつ大量に出力されてしまい、商品紹介文の改善として適切な出力とならない可能性がある。このため、予め評価視点表現を評価視点グループへと対応付けることで、この問題を回避する。

以上を踏まえ、本研究課題に対する枠組みの提案を図 3.1 に示す。まず、全レビュー文書集合を用いて、評価視点表現を評価視点グループの小規模な対応付け (辞書シード) に基づき、その他の評価視点表現を評価視点グループへ対応付ける (3.4.3 節)。以下、ここで得られる評価視点表現-評価視点グループの対応付けを評価視点グループ辞書、あるいは単に辞書と呼ぶ。次に、ある対象商品におけるレビュー集合から、辞書を用いて各評価視点グループの言及頻度を数え上げ、他商品の言及頻度との比較を通じて評価視点グループのランキングを行う (3.4.4 節)。最後に、最も高順位であった評価視点グループの上位 N 個を、説明文に相応しい内容として書き手に提示する。

以下、最初に本研究で用いる評価視点グループの定義を 3.4.2 節で示す。続いて、評価視点表現と評価視点グループの対応付けを 3.4.3 節、評価視点グループのランキングを 3.4.4 節で順次説明する。

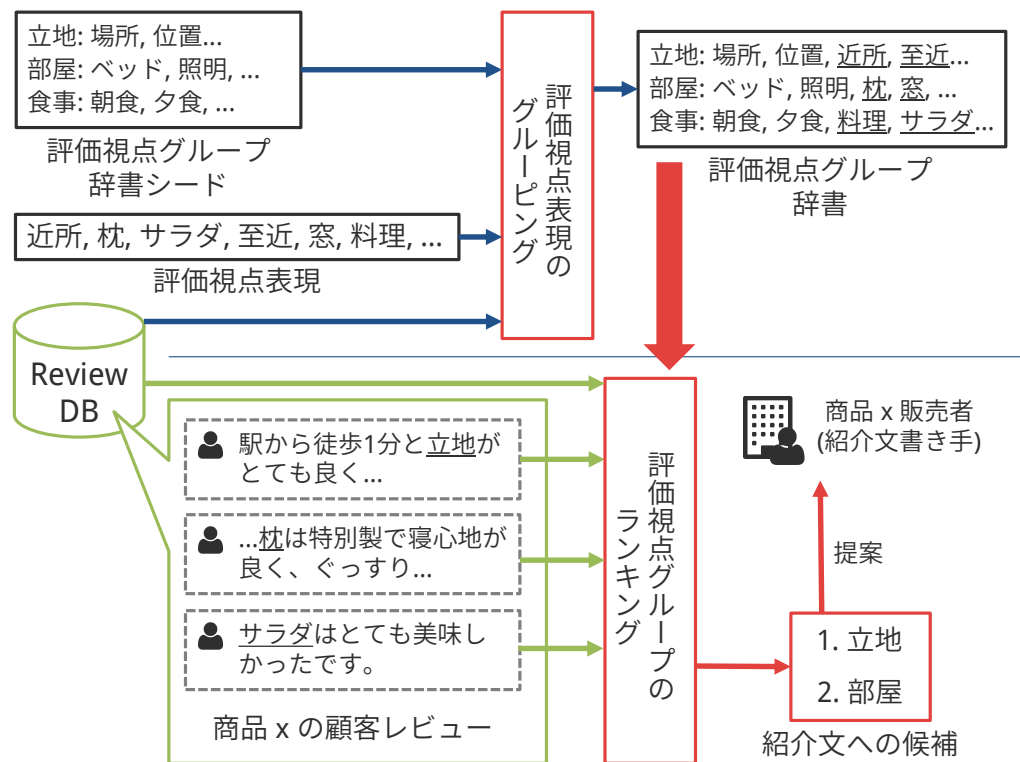


図 3.1: 提案手法の枠組み

3.4.2 評価視点グループの定義

本節では、本研究で用いる評価視点グループの定義を示す。どのような評価視点グループを用いるかは、対象とする商品と目的によって異なることを3.2.2節で述べた。この観点から考えると、本研究の商品は宿泊サービス・用途は商品紹介文改善のための情報であることから、評価視点グループには宿泊サービスの評価視点が意味的なまとまりにより分けられているものが望ましい。

ここで、用いる評価視点グループを考えるにあたって、評価視点グループの粒度について考える。評価視点グループの粒度とは、例えば宿泊サービスの「風呂」は「風呂のシャワー」「風呂の浴槽」などを包含し、「風呂のシャワー」は「シャワーの出」「シャワーの温度」などを包含していると言った階層関係において、どの階層を用いるかに相当する。用いる評価視点グループの粒度は、次の条件を満たしていることが望ましい。1) まず、本研究課題では評価視点グループ間の比較を通じて相応しい順に提示するため、商品紹介文の改善情報として同じ情報の有用性を持つことが望まれる。2) また、提案手法では他施設との比較を通じて評価視点グループをスコア付けするので、他施設との比較が可能な粒度が望まれる。

このような条件を満たせる粒度を考えるために、具体的に立地・風呂・食事を構成する評価視点として「駅からの距離」「露天風呂」「夕食のワイン」を考える。これに対し、条件1を満たすような評価視点グループとなるよう、相異なる評価視点の粒度を合わせることは困難である。すなわち、「駅からの距離」「露天風呂」「夕食のワイン」が同程度の重要性を持つかは、主観を排除して判断することは難しく、各宿泊施設の態様によっても異なる。また、条件2について考えると、「駅が近くにあるか」「大浴

場があるか・露天風呂があるか」「夕食を提供しているか・夕食にワインを提供しているか」と言った下位レベルの評価視点が存在するかは、宿泊施設によって異なる。

この問題は、どのような評価視点グループの粒度においても避けられないが、本研究では恣意性を可能な限り少なくし、多くの宿泊施設が共通に持つ観点から、最も粒度が大きい評価視点グループへの分割を行うこととした。以上の観点より、実際に「楽天トラベル」における商品・商品紹介文・レビュー文書の分析を通じて検討を行い、評価視点グループを最も基本的な 12 種類に分割した。評価視点グループの詳細を表 3.1 に示す。

3.4.3 評価視点表現の評価視点グループへの対応付け

評価視点表現の評価視点グループへの対応付けには、3.3.1 節で説明した Zhai et al. (2010) の手法を用いる。対応付けに半教師あり学習を用いる理由は、3.4.2 節で説明したような意味的レベルが高い評価視点グループは多数の評価視点表現を内包するため、人手による対応付けのコストが高いためである。半教師あり学習による手法では、小規模の辞書データから、全ての評価視点表現が対応付いた辞書を自動で獲得することができる。

3.3.1 節で説明した手法は、評価視点表現を評価視点に対応付ける手法である。一方で、本研究では評価視点の代わりに評価視点グループへ対応付ける、すなわち定義 2 で言う「評価視点グループを構成する表現」を推定する必要がある。そこで、本研究では用いるラベル集合を、評価視点ではなく評価視点グループとすることで、評価視点グループへの対応付けを実現する。

例えば、評価視点表現「チェックイン対応」「チェックアウト対応」「ルームサービスの対応」を考える。評価視点のレベルで考えると、評価視点表現「チェックイン対応」「チェックアウト対応」は評価視点「フロントの対応」、評価視点表現「ルームサービスの対応」は評価視点「ルームサービス対応」などに分類できる。これに対し、評価視点グループで考える場合、いずれの評価視点表現とも評価視点グループ「サービス」が対応付いていると考える。

3.4.4 評価視点グループのランキング

評価視点グループのランキングには、3.3.2 節で説明した対数尤度比 (Neyman & Pearson, 1928) を用いる。対数尤度比を二群のテキスト集合における事象の特徴度合いの指標として用いる利点を 3.3.2 節で述べたが、評価視点グループのランキングに適用する場合も同様の効果がある。すなわち、評価視点グループの特徴度合いを、注目商品のレビュー文書集合における頻度だけでなく、他商品のレビュー文書集合における頻度に基づいて測ることができる。これにより、他商品レビューでは言及が少ないが対象商品レビューでは言及が多いような評価視点グループに高いスコアを与えることができる。

従来研究において、対数尤度比は単語のコロケーション抽出 (Manning & Schütze, 1999) や特徴語抽出 (内山他, 2004; 小林・小木曾, 2013)・特徴語ランキング (乾他, 2013) を用いられていることを 3.3.2 節で述べた。これらの研究は、いずれも単語レベルでのスコア付けを目的としている。これに対し、本研究では単語ではなく評価視点グループにスコアを与える必要がある。そこで、本節では評価視点グループを構成する評価視点表現の出現回数に基づいた値を利用することで、評価視点グループのスコアを与える方法を示す。

具体的に、商品 h の評価視点グループ g について、対数尤度比によるスコア付け手法を示す。まず、

表 3.1: 本研究で用いる宿泊サービスを対象とした評価視点グループの一覧

評価視点グループ	説明
サービス	<p>スタッフ・従業員の接客態度、おもてなし。</p> <ul style="list-style-type: none"> 「ウェルカムドリンクのサービス」「無料送迎サービス」「入浴剤のサービス」などは含まない。それぞれ<食事>、<立地>、<アメニティ>に含める。
立地	<p>施設への/施設からのアクセス性。</p> <ul style="list-style-type: none"> 「駐車場」「送迎サービス」についても含む。 「コンビニへ近い」「アウトレットモールに近い」「羽田空港へ近い」など、周辺施設や機関についての記述も含む。 「湖畔の側に立地…」のように、利便性ではなく景観が重視されている場合は<景観>に含める。但し、例えば湖畔のアクセス性が高く、立地のアピールとなる場合は<立地>に含める。
景観	<p>施設からの眺め（海、山、夜景など）、雰囲気。</p> <ul style="list-style-type: none"> 施設と直接関係のない近隣施設・機関からの景観は含めない。
外装・内装	<p>施設の外観、フロント・ロビー・廊下など館内の見た目や雰囲気。</p> <ul style="list-style-type: none"> 宿泊部屋については<部屋>に含める。
部屋	<p>部屋の広さ、雰囲気、清掃。</p> <ul style="list-style-type: none"> 電灯、ベッド（布団・枕）についても含む。 付属の風呂については、<風呂>に含める。バス・トイレ一体型などの場合、トイレなど風呂以外の記述を対象とする。
室内設備	<p>部屋内の設備、備品、機器。</p> <ul style="list-style-type: none"> テレビ、冷蔵庫、エアコン、有線・無線 LAN などを含む。 「貸し出しの加湿器」など、貸与機器についても含む。
館内設備	<p>施設が営業・管理する売店、自販機、コインランドリー、卓球場、ゲームコーナーなど、館内の設備・施設。</p> <ul style="list-style-type: none"> ビジネスホテルなどでコンビニが併設されている場合、施設が直接営業していないので<立地>に含める。
アメニティ	<p>タオル、石鹸、シャンプー、歯磨きセットなどの使い捨て備品。</p>
風呂	<p>室内バス、大浴場、露天風呂など、施設が提供する風呂。</p> <ul style="list-style-type: none"> （露天）風呂付客室の風呂についての記述は<部屋>ではなく<風呂>に含める。
食事	<p>施設が提供する食事。</p> <ul style="list-style-type: none"> 「部屋食」のような食事についての付加要素についても含む。 食事会場のスタッフの行動については、食事に関わる割合が大きい場合（バイキングの補充など）<食事>へ、接客に関わる割合が大きい場合（配膳態度など）<サービス>に含む。 食事会場が館外（隣接のレストラン等）で提供されていても、多くの宿泊プランで一体になっている場合は<食事>に含む。
価格	<p>宿泊費用。</p> <ul style="list-style-type: none"> 「ポイント還元」などのポイントサービスについても含む。
ウェブ	<p>ウェブサイトの作り・案内、ユーザからの評価・評判。</p> <ul style="list-style-type: none"> 「〇〇アワード受賞」「ユーザ評価 4.5」なども含む。

3.3.2 節における文書 A, B はそれぞれ以下のように置く。

A : 商品 h のレビュー文書集合

B : 商品 h 以外の全商品のレビュー文書集合

ここで、二群のレビュー文書集合における表現（単語またはフレーズ）を順次見た際、観測された表現 v について確率変数 X, Y を以下のように置く。

$$X = \begin{cases} 1 & \text{if 表現 } v \text{ は評価視点グループ } g \text{ を構成する表現である} \\ 0 & \text{otherwise} \end{cases} \quad (3.31)$$

$$Y = \begin{cases} 1 & \text{if 表現 } v \text{ は商品 } h \text{ のレビュー文書集合の表現である} \\ 0 & \text{if 表現 } v \text{ は商品 } h \text{ 以外の全商品のレビュー文書集合の表現である} \end{cases} \quad (3.32)$$

この場合、確率変数 X, Y が独立であるということは (H_{indep})、すなわち g を構成する表現の出現確率が、注目商品のレビュー文書集合・他商品のレビュー文書集合で同じとすることを意味する。逆に、従属であるということは (H_{dep})、 g を構成する表現の出現確率が異なることを意味する。これらの仮説の下でのデータ全体の出現確率の比を式 (3.25) で求めることにより、他商品レビュー文書集合と比較して商品 h のレビュー文書集合において言及回数が多い評価視点グループ g に高いスコアを与えることができる。

以上の式 (3.31), (3.32) に基いて、対数尤度比に用いる頻度値 a, b, c, d を置き換えると以下の通りである。

a = 商品 h のレビュー文書集合における、評価視点グループ g を構成する表現の出現回数

b = 全商品のレビュー文書集合における、評価視点グループ g を構成する表現の出現回数

c = 商品 h のレビュー文書集合における、評価視点グループ g に属さない全表現の出現回数

d = 全商品のレビュー文書集合における、評価視点グループ g に属さない全表現の出現回数

以上の置き換えを行うことで、式 (3.30) の値は、評価視点グループ g が商品 h のレビュー文書集合で特徴的なほど高い値を取り、商品 h 以外のレビュー文書集合で特徴的なほど低い値を取るようになる。

一方で、上記で a, b, c, d に用いている表現の頻度情報はレビュー文書によっては偏りが大きくなる恐れがある。すなわち、ある顧客が特定の評価視点グループに偏って長文のレビュー文書を作成している場合、特に a, c の値に影響を与える可能性がある。そこで、「評価視点グループ g を構成する全表現の出現回数」を、「構成する表現を含むレビュー文の出現回数」または「構成する表現を含むレビュー文書の出現回数」と読み替えたバリエーションを考える。

W-LLR: 評価視点グループ g を構成する表現の出現回数

S-LLR: 評価視点グループ g を構成する表現を含むレビュー文の出現回数

R-LLR: 評価視点グループ g を構成する表現を含むレビュー文書の出現回数

これらの各手法を、それぞれ W-LLR, S-LLR, R-LLR と呼ぶ。最適な設定は評価実験を通して考察する。

加えて、本研究の目的である商品紹介文に相応しい評価視点グループの提示に合わせ、上記の値をポジティブな文脈に制限する。すなわち、商品への満足を述べているなどのポジティブな文脈で言及が多い評価視点グループを優先的に提示し、苦情などのネガティブな文脈で言及が多い評価視点グループを

無視したい。そこで、各レビュー文の極性判定を行う前処理を加える*2。具体的な手順は以下の通りである。まず、レビュー文がポジティブか否かを判定する二値分類器を予め作成しておく。次に、対数尤度比の計算を行う前に、全レビュー文書集合から二値分類器によりポジティブ文のみを抽出する。この場合の各変数 a, b, c, d は以下の通り対応する (W-LLR の場合)。

- a = 商品 h のレビュー文書集合のうち、ポジティブ文において g を構成する全表現の出現回数
- b = 全商品のレビュー文書集合における、ポジティブ文において g を構成する全表現の出現回数
- c = 商品 h のレビュー文書集合における、ポジティブ文において g に属さない全表現の出現回数
- d = 全商品のレビュー文書集合における、ポジティブ文において g に属さない全表現の出現回数

以上の値から対数尤度比の計算を行うことで、評価極性を考慮した評価視点グループのスコア付けが実現できる。

3.5 評価実験

3.5.1 実験条件

実験に用いるデータには、楽天トラベルにおける宿泊施設の紹介文及びレビュー集合を用いた。紹介文のデータとして宿泊施設の詳細情報ページに掲載されているテキストを、レビュー集合のデータとして「楽天データ公開*3」に収録されているデータを用いた。このうち、レビュー件数が10件から1000件の施設を選択し、13,664施設、2,254,307レビューを得た。

レビュー文書の前処理は次のように行った。まず、レビュー文書の文分割は句点などを文の分割点とするなどの簡易的なヒューリスティック (詳しくは第A章で説明する) によって行った。また、単語分割には MeCab 0.996 (Kudo et al., 2004) を利用した。単語分割後、複合名詞と「<名詞>の<名詞>」表現については1つのフレーズとして結合した。

評価視点グループ辞書は、以下のように構築した。まず、名詞及び名詞を含むフレーズを、全施設のうち1%以上の施設で用いられている9,844表現を評価視点表現として抽出した。次に、抽出された表現のうち、最も使用頻度が高い281表現 (上位約3%程度) に対し、評価視点グループを対応付けた。最後に、3.4.3節で説明した手法により、全ての評価視点表現に対して評価視点グループを対応付け、評価視点グループ辞書とした。

レビュー文がポジティブか否かを判定する二値分類器の作成には SVM (Boser et al., 1992; Cortes & Vapnik, 1995; Vapnik & Lerner, 1963) を用いた。SVMの実装には scikit-learn 0.15.2 (Pedregosa et al., 2011) を、学習データには楽天データ公開に含まれる筑波大学文単位評価極性タグ付きコーパス (TSUKUBA コーパス) を利用した。分類の素性には、レビュー文を構成する単語を用いた。単語分割には MeCab 0.996 (Kudo et al., 2004) を利用した。文のラベルは、コーパス中で「褒め」ラベルが付与されている文をポジティブ文、それ以外のラベルが付与されている文を非ポジティブ文とした。以上のデータを用い、表に3.2に示すパラメータにより分類器を作成した。なお、表3.2のパラメータは、パラメータ探索範囲で示す組み合わせからグリッドサーチを行って設定したパラメータである。予備実験

*2 なお、予備実験として、ポジティブ文の抽出処理を行わない場合についての性能も調査した。その結果、性能は表3.4, 3.5に示す性能よりも僅かに低下した (対数尤度比による手法では、1-bestの適合率・再現率が2-3ポイント低下するなど)。性能に大きな差が見られなかった主な理由としては、楽天トラベルにおけるレビュー文の多くがポジティブ文である点がある。

*3 楽天データ公開: <http://rit.rakuten.co.jp/rdr/>。データは2012年当時の公開データを用いた。

表 3.2: ポジティブ文分類器の設定

項目	設定値
SVM	scikit-learn 0.15.2
カーネル	RBF
ペナルティ項 C	3.60
RBF 係数 γ	-5.20
パラメータ探索範囲	
カーネル	{ Linear, RBF }
ペナルティ項 C	$2^{\{-5...15\}}$
RBF 係数 γ	$2^{\{-15...3\}}$
交差検定の分割数	5
スコア基準	「褒め」分類性能 F 値の最大化

において、5 分割交差検定による評価では適合率 83.08%、再現率 79.79% を確認しており、本研究で用いるために十分な性能であった。

評価視点グループのランキング手法として、以下の手法の比較を行う。まず、3.4.4 節で説明した対数尤度比に基づく手法 LLR を用いる。これに対し、評価視点グループの頻度を数える単位を単語 W-LLR, 文 S-LLR, 文書 R-LLR と変化させて比較する。加えて、ランキング手法のバリエーションとして、単純な頻度でランキングする手法、T-スコアでランキングする手法とも比較する。それぞれについて、頻度を数える単位を単語・文・文書と変化させた手法を W-TF, S-TF, R-TF 及び W-Tscore, S-Tscore, R-Tscore とする。

評価視点グループ対応付けの評価は、次のようにして行う。まず、ランダムサンプリングにより 100 レビュー文書を抽出し、レビュー文書に含まれるレビュー文 450 文に対し、各レビュー文中で言及されている評価視点グループをタグ付けしたデータを作成した（一つのレビュー文に対し、評価視点グループが対応しない場合や評価視点グループが複数対応する場合がある）。次に、これらのレビュー文に対し、各レビュー文から評価視点グループ辞書に含まれる単語・フレーズを抽出し、単語・フレーズと対応する評価視点グループを各文のラベルとみなし、正解データとの比較を行った。文レベルで評価を行う理由は、単語・フレーズ単位での評価は、システム出力と正解データのフレーズ境界が必ずしも一致しないため正解の判定が困難なためである。評価指標としては、各評価視点グループごとの適合率 (Precision) ・再現率 (Recall) を用いており、具体的な計算式は以下の通りである。

$$\text{Recall} = \frac{\text{注目評価視点グループにおいて正しく推定できた文数}}{\text{評価データ中における注目評価視点グループの文数}}$$

$$\text{Precision} = \frac{\text{注目評価視点グループにおいて正しく推定できた文数}}{\text{システムが注目評価視点グループとして出力した文数}}$$

評価視点グループのランキングの正解データとして、次の 2 つのデータを用意した。

- **As-is セット**: 実際に用いられている商品紹介文を正解と考え、各商品紹介文で言及されている評価視点グループを正解データとする。
- **Revise セット**: 人手により、商品紹介文に相応しいと考えられる評価視点グループを付与したデータを正解データとする。

まず、As-is セットについては、ランダムサンプリングによって得られた 126 施設のうち、商品紹介文として適切ではない*4と判断した施設を排除した 116 施設を選択し、各商品紹介文で言及されている評価視点グループをタグ付けした。また、Revise セットについては、ランダムに得られた 126 施設に対し、商品紹介文として相応しい順に評価視点グループをタグ付けした。タグ付けに用いた情報源は、現在の商品紹介文、施設の紹介ページ、直近のレビュー文書 50 件である。

評価視点グループのランキングの評価には、次の評価指標を用いる。まず、As-is セットには正解データに対して順位が与えられていないため、システムが出力する N-best の n の適合率・再現率、及び平均適合率 (Average Precision) を評価指標とした。適合率・再現率の具体的な計算式は以下の通りである。

$$\text{Recall} = \frac{\text{注目施設において正しく推定できた評価視点グループ数}}{\text{注目施設に含まれる評価視点グループ数}}$$

$$\text{Precision} = \frac{\text{注目施設において正しく推定できた評価視点グループ数}}{\text{システムが出力した評価視点グループ数 (n)}}$$

また、平均適合率は、適合する評価視点グループ数が得られた n における適合率の平均である。

一方、Revise セットには正解データには順位情報が含まれるため、評価指標としては順位付けを考慮できる NDCG (Normalized Discounted Cumulative Gain) (Järvelin & Kekäläinen, 2002) を用いる。NDCG はランキング手法の出力と正解データの類似度に基づく指標であり、以下の式で計算する。

$$\text{NDCG}@n = \frac{\text{DCG}@n}{\text{IDCG}@n} \quad (3.33)$$

$$\text{DCG}@n = \text{rel}_1 + \sum_{i=2}^n \frac{\text{rel}_i}{\log_2 i} \quad (3.34)$$

ここで、 n は N-best の設定、 rel_i は i 番目として提示された評価視点グループの適合度 (relevance grade)、IDCG@ n は正規化項である。本研究では適合度 rel_i の評価関数として、以下に示す関数を用いた。

$$\text{rel}_i = \frac{1}{\log(1+r)} \quad (3.35)$$

ここで、 r は i 番目として提示された評価視点グループの、正解データにおける実際の順位である。式 (3.35) は NDCG の適合度評価関数として広く用いられており、Wang et al. (2013) によってランキング手法間の性能差を測るために適した NDCG の適合度評価関数の一つであることが示されている。

3.5.2 実験結果

評価視点表現の評価視点グループへの対応付け

まず、評価視点表現の評価視点グループへの対応付けの実験結果を示す。表 3.3 に、各評価視点グループごとの対応付けの性能を示す。

表 3.3 より、評価視点グループのランキングのための前処理と言う観点からは概ね十分な性能が得られているが、特に <室内設備> と <ウェブサイト> の性能が低いことが分かる。これは、レビュー文書において室内設備やウェブなどに関する記述が少ないため、分類が困難になっていると言う原因がある。筆者の観察によると、楽天トラベルのレビュー文書においては <室内設備> と <ウェブサイト> などの

*4 他の商品紹介文と比較して極端に短い紹介文や、商品の紹介を行っていない (例えば、ポイントや予約に関する制約事項など、商品紹介文とは異なる場所で説明すべき事項が書かれている) 紹介文が含まれる。

表 3.3: 評価視点表現の評価視点グループへの対応付けの性能

評価視点グループ	適合率	再現率
サービス	71.67%	70.49%
立地	64.86%	70.59%
景観	55.56%	38.46%
内装・外装	85.71%	30.77%
部屋	78.18%	64.18%
室内設備	52.17%	36.36%
館内設備	28.57%	8.70%
アメニティ	83.33%	31.25%
風呂	88.64%	88.64%
食事	71.62%	79.10%
価格	76.92%	44.44%
ウェブ	36.36%	12.90%

表 3.4: As-is セットでの設定における評価視点グループのランキングの性能

手法	平均 適合率	1-best			2-best			3-best		
		適合率	再現率	F 値	適合率	再現率	F 値	適合率	再現率	F 値
R-LLR	0.666	0.643	0.336	0.442	0.487	0.509	0.498	0.397	0.623	0.485
S-LLR	0.663	0.635	0.332	0.436	0.474	0.495	0.484	0.394	0.618	0.481
W-LLR	0.658	0.626	0.327	0.430	0.465	0.486	0.476	0.397	0.623	0.485
R-Tscore	0.613	0.522	0.273	0.358	0.417	0.436	0.427	0.374	0.586	0.457
S-Tscore	0.623	0.557	0.291	0.382	0.400	0.418	0.409	0.374	0.586	0.457
W-Tscore	0.615	0.530	0.277	0.364	0.417	0.436	0.427	0.374	0.586	0.457
R-TF	0.596	0.496	0.259	0.340	0.422	0.441	0.431	0.359	0.564	0.439
S-TF	0.596	0.496	0.259	0.340	0.430	0.450	0.440	0.362	0.568	0.442
W-TF	0.587	0.478	0.250	0.328	0.417	0.436	0.427	0.354	0.555	0.432

評価視点グループに関しては顧客が何か特別な特徴を見出していない限り書かない傾向があることが分かっている。このような評価視点グループへの対応付けの性能が評価視点グループのランキングに与える影響については、次節で説明する。

評価視点グループのランキング

続いて、評価視点グループのランキングの実験結果を示す。As-is セット・Revise セットにおける実験結果をそれぞれ表 3.4, 3.5 に示す。表 3.5 の @ n は N-best の設定を示す。

最初に、評価視点グループ頻度の数え方を固定し、評価視点グループのランキング手法について比較する (R-LLR, R-Tscore, R-TF)。表 3.4, 3.5 を見ると、対数尤度比を用いた手法 R-LLR がいずれの実験設定においても高い性能であることが分かる。同様の傾向は、S-LLR, S-Tscore, S-TF 間および W-LLR, W-Tscore, W-TF 間でも観察できる。この結果から、本研究課題においては、対数尤度比に基づくランキ

表 3.5: Revise セットでの設定における評価視点グループのランキングの性能

手法	NDCG				
	@1	@2	@3	@4	@5
R-LLR	0.743	0.736	0.737	0.752	0.770
S-LLR	0.751	0.741	0.739	0.745	0.764
W-LLR	0.777	0.735	0.736	0.737	0.760
R-Tscore	0.567	0.597	0.639	0.663	0.693
S-Tscore	0.599	0.607	0.647	0.673	0.701
W-Tscore	0.589	0.614	0.643	0.657	0.690
R-TF	0.674	0.663	0.676	0.707	0.753
S-TF	0.695	0.678	0.682	0.718	0.760
W-TF	0.648	0.649	0.654	0.705	0.751

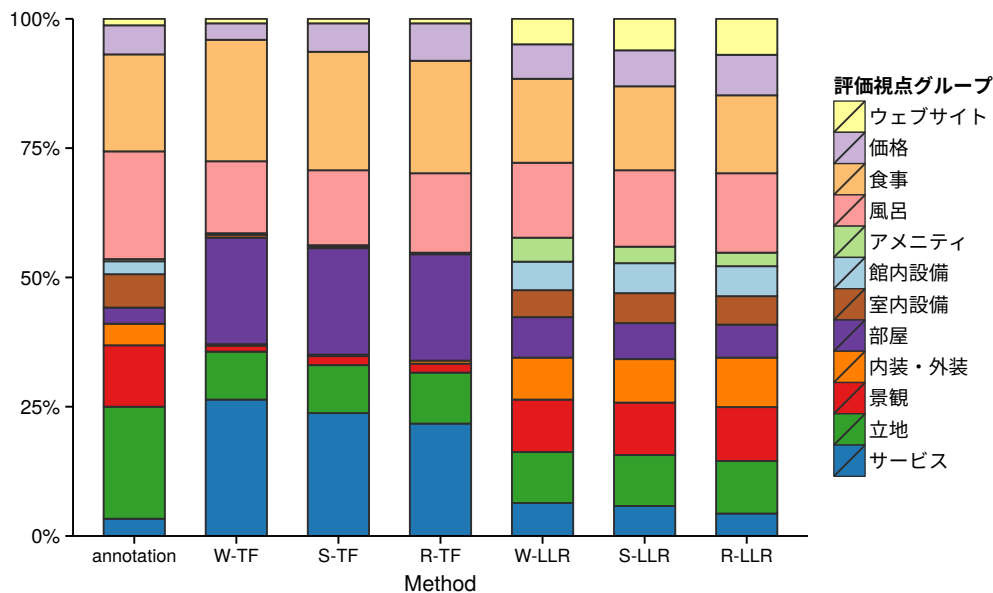


図 3.2: 3-best 設定における Revise セットと各ランキング手法による出力結果の分布

ングが最も適切であることが分かった。次に、評価視点グループのランキング手法を固定し、評価視点グループ頻度の数え方の間で比較する (R-LLR, S-LLR, W-LLR)。この場合は、表 3.4 における 1-best, 2-best 及び平均適合率、表 3.5 における NDCG@2 から NDCG@5 については R-LLR, S-LLR が W-LLR を上回っているものの、差はわずかである。同様の傾向は R-Tscore, S-Tscore, W-Tscore 及び R-TF, S-TF, W-TF の間でも観察できた。この結果から、本研究課題においては、評価視点グループ頻度の数え方はランキング手法よりもパラメータの影響としては弱いと言える。

ランキング手法による出力結果の違いを分析するため、3-best 設定における各ランキングの出力結果の分布を図 3.2 に示す。頻度による手法 W-TF, S-TF, R-TF の出力は、各評価視点グループに含まれる評価視点表現がレビュー文書で出現した回数を意味するが、正解データとは特に <サービス>、<部屋>、<

表 3.6: S-LLR による出力例。システム出力は S-LLR のスコア順に並んでいる。

施設	紹介文/ システム出力	評価視点 グループ	紹介文/ システム出力の評価視点グループに対応するレビュー文例
#1	システム	(立地)	日本海と最上川が会おう港町
		風呂	露天風呂、打たせ風呂、熱め、ぬるめ、いろいろな種類のお風呂を楽しめました。
		食事	食事も海と山の新鮮な物を頂き大変美味しかったです。
		部屋	お部屋がとてもきれいでリニューアルしたてのようでした。
#2	システム	(外装・内装)	9月10日(土)リニューアル外装・内装オープン!
		風呂	家族風呂も4つあり、時間も24時間入ってOKだったので、とってもいいお湯に何回も入ることができて大満足です。
		食事	生ビールやコーヒーの飲み放題にびっくり、お部屋もきれいでしたし、食事も個室で美味しかったです。
		アメニティ	タオルも使い放題、アメニティの他、化粧品や乳液もあり、本当に感激しました。

景観>、<館内設備> について大きく異なっている。まず、<サービス>、<部屋> については、楽天トラベルのレビュー文書においては、特別な事情や体験がなくても多くの顧客が書くと言う特徴を持つ。一方、<景観>、<館内設備> については、何らかの特別な体験がない限り触れないと言う特徴を持つ。このため、単純な言及数のみで判定すると、このようなバイアスに影響されてしまうと言う問題がある。加えて、<館内設備> については、表 3.3 に示すように再現率が低いことから、言及されているにも関わらずカウントできていないレビュー文が存在すると言う原因がある。一方で、対数尤度比による手法 W-LLR, S-LLR, W-LLR による出力は、より正解データに近い分布を持つ。これは、対数尤度比によれば他施設のレビュー文書集合との比較を通じ、例えば他施設でも言及数が多い<サービス>・<部屋> 等に対しては低いスコアに、他施設では言及数が少ない<景観>・<館内設備> 等に対しては高いスコアになるよう重み付けを行えているためである。この重み付けは、頻度による手法が<館内設備> の評価視点グループ辞書による影響を受けていた問題を軽減できるという効果もある。他方で、対数尤度比による手法では<ウェブサイト> が過大評価され、<立地> が過小評価されていると言う特徴がある。これは、評価視点グループの提示にあたってレビュー文書情報のみを用いており、商品紹介文に関わる情報が考慮されていないことによる。性能の改善にあたっては、商品紹介文に関わる事前知識、すなわちどのような評価視点グループが商品紹介文として好まれるかと言った情報を考慮する必要があると考える。

最後に、3-best において最も性能が高かった S-LLR による出力例を表 3.6 に示す。施設番号 #1, #2 いずれに対しても、紹介文には書かれていない評価視点グループを提示しているが、レビュー文例を見ると紹介文に追加するのに相応しい評価視点グループであると言える。この情報を用いることで、施設番号 #1, #2 の説明文では 1 つの評価視点グループのみにしか触れていないが、提示された評価視点グループである「風呂」や「食事」を追加すると言う判断が可能である。また、施設番号 #2 については、リニューアルから時間が経過し、訴求する評価視点グループを変更したい際に、提示された評価視点グループを用いることも可能である。

3.6 関連研究

ここで、本研究と関連する研究について述べる。まず、本研究と類似する対象の構想過程における文書作成支援を目的とした研究について述べる。これらの研究概要は2.1.2節で述べたが、ここでは特に本研究課題と関連が深い研究について、本研究との違いを手法の比較を中心に説明する。続いて、本研究課題と関連して、評価視点表現の構造化手法、評価視点のランキング手法について述べる。

本研究と最も関連が深い研究は、2.1.2節で挙げたキャッチコピー生成研究である。中野・鬼沢(2008)は商品説明文のキーワードと共通した概念をもちながら意外性のある語句を連想し、連想語句を用いたキャッチコピー生成手法を提案している。中野・鬼沢(2008)の手法では、連想語句の生成は笠原他(1994)の提案した概念ベースにおける概念語間の距離に依拠しているが、本研究では顧客レビューの分布に基づいて評価視点グループのスコア付けを行っている。

また、幅野・浦谷(2013, 2014)は、映画のあらすじとレビュー文書集合を用い、既存のキャッチコピーにおける文法構造を基に映画キャッチコピーを生成する手法を提案している。幅野・浦谷(2013, 2014)の手法では、レビュー文書は既存のキャッチコピーを置換するための単語を抽出するために用いられている。これに対し、本研究ではレビュー文書の利用方法、すなわち他商品も含めたレビュー文書を用いた評価視点グループのスコア付けへ用いている点が異なる。

更に、キャッチコピー生成のためにキーワードを要する手法も提案されている(Yamane & Hagiwara, 2015, 2013; 松平・萩原, 2005, 2004)がある。本研究は構想過程における支援が目的であることから、どのようにキーワードを入力すべきか未知の状態を前提としている点で異なるが、本研究の出力に基づいたキーワードを用いることで、これらの手法によるキャッチコピー生成へと繋げることができる。

評価視点表現の構造化を目的とした研究に注目すると、トピックモデルなどによる教師なし学習に基づく手法が提案されている(Brody & Elhadad, 2010; Chen et al., 2013a,c,b; Guo et al., 2009; Titov & McDonald, 2008a,b; Zhai et al., 2011)。しかし、これらの教師なしによる構造化手法は、必ずしも人間が解釈可能な単位の分類結果が得られるとは限らないため、本研究の目的である評価視点グループの提示による文書作成支援には不適である。一方、少量の事前知識を用いることで、事前知識を用いない場合よりも高性能かつ解釈が容易な構造化を実現する研究も行われている(Carenini et al., 2005; Zhai et al., 2010)。Carenini et al. (2005)の手法は、評価視点表現を評価視点に関連付ける際に WordNet の同義語情報を利用しているが、商品ドメインによっては同義語でなくても同じ評価視点と見なせる評価視点が存在し、特に評価視点グループのレベルではこの問題が顕著であることから、本研究に用いる処理としては相性が悪い。これに対し、Zhai et al. (2010)の手法は、同義語などの制約が不要であり、評価視点グループへの拡張が容易であることから、本研究において評価視点表現の評価視点グループへの対応付け手法の基礎として用いた。

一方で、評価視点をランキングする研究に注目すると、Yu et al. (2011); Zhang et al. (2010); 乾他(2013)の手法がある。このうち、Yu et al. (2011); Zhang et al. (2010)は与えられたドメイン全体における重要性と言う観点から評価視点のランキングを行っており、本研究の目的である商品別の評価視点のランキングとは目的が異なる。これに対し、乾他(2013)は商品別の評価視点表現のランキングを行うため、本研究と同様に対数尤度比に基づいた手法を提案している。本研究は評価視点表現ではなく評価視点グループのランキングを行っていることから、乾他(2013)が対数尤度比の応用として評価視点表現でのランキングを示したのに続いて、本研究では更に対数尤度比を評価視点グループのランキングへの

応用を示した研究であると言える。

3.7 まとめと今後の課題

本論文では、商品紹介文に相応しい評価視点グループを提示する課題を新たに提案した。また、本課題に対して、評価視点表現の評価視点グループへの対応付け・評価視点グループのランキングの組み合わせで行う枠組みを提案した。実際に評価視点グループを提示する実験を通じ、提案した枠組みの有効性を確認した。

今後の課題として、以下の項目を考えている。

まず、比較する商品を、特に競争の対象となりやすい商品に選択することを考えている。例えば、宿泊サービスであれば、対象とする宿泊施設に至近の宿泊施設が競争の対象となりやすい。この場合、比較の際に近隣宿泊施設のレビュー文書を用いることで、それらの施設と比較して特徴的な評価視点グループを提示できる。これにより、競争的商品との差別化ができる紹介文の作成支援が可能になる。

また、評価視点グループの提示に合わせ、根拠となるレビュー文・レビュー文書などの提示ができることが望ましい。これにより、提示された評価視点グループが顧客にとってどのように好評であったかが分かるため、紹介文の作成の際に役立つと考えている。多くの場合、関連する文・文書は多数存在するため、これらに対しても他商品のレビュー文書集合との比較を通じ、他商品では見られない特徴的な文・文書を提示したいと考えている。更にこの考えを推し進めると、ポジティブ感想以外（例えば苦情や要望・要求など）から評価視点グループをスコア付けをし、他施設と比較して苦情が多かった評価視点グループと、併せて根拠となるレビュー文・レビュー文書などを提示できるようになる。これらの情報は、商品販売者にとっては商品の改善に繋げるための重要な情報である。このような情報を抽出する課題は本研究の目的である商品紹介文の作成支援からは外れるが、これらの課題を含め、本研究で提案した枠組みの更なる応用可能性についても検討したい。

第4章

対をなす二文書間における文対応関係の推定

本研究課題では、対となる二つの文書、例えば顧客からのメール文書やレビュー文書と販売者の応答文書などを対象とし、文レベルでの対応関係を推定する課題の提案と解決手法の検討を行う。本章では、最初に研究課題の背景について説明し、研究課題の提案を行う。次に、文対応の定義を行い、レビュー文書と応答文書を対象としたデータの分析結果を示す。続いて、提案手法の説明に必要な基盤技術と諸概念を導入し、これらを利用して提案手法の説明を行う。更に、評価実験を通じて、提案手法の有効性を示す。最後に、本研究の目的と関連した研究の概要と、本研究との違いについて述べる。

4.1 研究課題の背景と概要

今日までに、人間による言語使用の仕組みを解明する試みが単語・文・発話・文書など様々な単位に注目して行われて来た。特に、これらの種類や相互関係（例えば単語であれば品詞や係り受け関係、文であれば文役割や修辭構造など）にどのようなものがあるか、どのように利用されているかを明らかにする研究が精力的になされて来た。

例えば、ある一文書内に登場する節と言う単位に注目すると、主な研究として Mann & Thompson (1987); Mann et al. (1992) による修辭構造理論 (Rhetorical Structure Theory; RST) がある。修辭構造理論では文書中の各節が核 (nucleus) と衛星 (satellite) の二種類に分類できるとし、さらに核と衛星の間にみられる関係を 21 種類に、核と核の間にみられる関係 (多核関係) を 3 種類に分類している。このような分類を用いて、節同士の関係を自動推定する研究も古くから行われている (Marcu, 1997; 田村・和田, 1998)。さらに、推定した関係を別タスクに利用する研究も盛んに行われている (Marcu, 1999; Marcu et al., 2000; Tu et al., 2013; 平尾他, 2013; 比留間他, 1999)。例えば、Marcu (1999); 平尾他 (2013); 比留間他 (1999) は、節の種類や節同士の関係を手がかりに重要と考えられる文のみを選択することで自動要約への応用を示している。また、Marcu et al. (2000); Tu et al. (2013) は、機械翻訳においてこれらの情報を考慮することで性能向上を実現している。

一方、本研究課題では従来研究の主な対象であった一文書や対話ではなく、ある文書 (往信文書) とそれに呼応して書かれた文書 (返信文書) の対を対象とし、往信文書中のある文と返信文書中のある文との間における文レベルでの呼応関係 (以下、**文対応**と呼ぶ) に注目する。なお、本研究課題において、文書対は異なる書き手によって書かれたものとする。このような文書対の例として「電子メールと返信」、「電子掲示板の投稿と返信」、「ブログコメントの投稿と返信」、「質問応答ウェブサイトの質問投稿

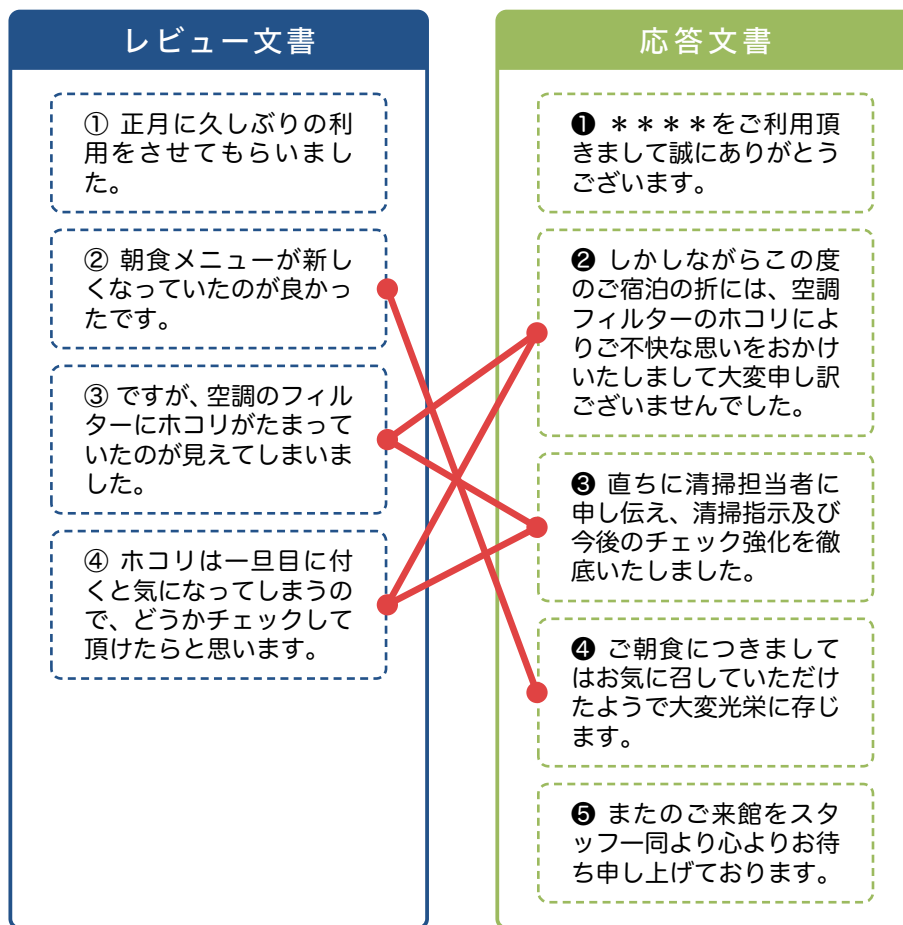


図 4.1: レビュー文書 - 応答文書対における文対応の例。文番号は、説明の都合上筆者が付与した。

と応答投稿」などがあるが、本研究課題では特に「商品やサービスに対するレビュー文書とサービス提供者の返信文書」を対象にする。レビュー文書と応答文書における文対応の例を図 4.1 に示す。図中の文同士を結ぶ直線が文対応を示しており、例えば返信文「④ ご朝食につきましてはお気に召していただけただようで大変光栄に存じます。」は返信文（レビュー文）「② 朝食メニューが新しくなっていたのが良かったです。」を受けて書かれた文である。

本研究課題では、文書レベルで返信・返信の対応が予め分かっている文書対を入力とし、以上に述べたような文対応を自動で推定する課題を新たに提案し、解決方法について検討する。これら文書対における文対応の自動推定が実現すれば、様々な応用が期待できる点で有用である。具体的な応用例について、レビュー文書・返信文書対における文対応推定の場合を中心に説明する。

- (1) **未対応文の検出による返信文書作成の支援:** 返信文書と返信文書を入力して自動で文対応を特定できるということは、逆に考えると返信文書の中で対応が存在しない文が発見できることでもある。この推定結果を利用し、ユーザが返信文書を作成している際に「返信文書中の対応がない文」を提示することで、返信すべき事項に漏れがないかを確認できる文書作成支援システムが実現できる。このシステムは、レビュー文書・返信文書対に適用した場合は顧客への質問・クレームへの対応支援に活用できる他、例えば質問応答サイトのデータに適用した場合は応答作成支援など

にも利用できる。

- (2) **定型的返信文の自動生成:** (1) の考えを更に推し進めると、文対応を大量に収集したデータを用いることで、将来的には定型的な返信文の自動生成が可能になると期待できる。大規模な文対応データを利用した自動生成手法は、例えば Ritter et al. (2011); 長谷川他 (2013) が提案しているが、いずれも文対応が既知のデータ（これらの研究の場合はマイクロブログの投稿と返信）の存在が前提である。しかし、実際には文対応が既知のデータは限られており、未知のデータに対して自動生成が可能となるだけの分量を手でタグ付けするのは非常に高いコストを要する。これに対し、本研究が完成すればレビュー文書・返答文書対をはじめとした文対応が未知のデータに対しても自動で文対応を付与できるため、先に挙げた様々な文書において往信文からの定型的な返信文の自動生成システムが実現できる。定型的な返信文には、挨拶などに加え、同一の書き手が過去に類似した質問や要望に対して繰り返し同様の返信をしている場合などが含まれる。
- (3) **非定形的返信文の返答例提示:** (2) の手法の場合、自動生成できるのは定型的な文に限られる。一方、例えば要望や苦情などの個別案件に対する返答文作成の支援は、完全な自動生成の代わりに複数の返答例を提示することで実現できると考えている。これを実現する方法として、現在返答しようとしている往信文に類似した往信文を文書対のデータベースから検索し、類似往信文と対応している返信文を複数提示する手法がある。返信文の書き手は、返答文例の中から書き手の方針と合致したものを利用ないし参考にすることで返信文作成の労力を削減できる。
- (4) **文書対群の情報整理:** 複数の文書対から、文対応が存在する文対のみを抽出することでこれら文書対の情報整理が可能になる。例えば、「このサービス提供者は（または要望、苦情など）に対してこのように対応しています」と言った一覧を提示できる。これを更に応用すれば、将来的には FAQ の（半）自動生成や、要望・苦情への対応率・対応傾向の提示などへ繋がれると考えている。

一方で、文書対における文対応の自動推定課題は以下のような特徴を持つ。なお、これらの特徴が実際のデータで観察されることは、後に 4.2.3 節で詳しく述べる。

- (1) **対応する文同士は必ずしも類似しない:** 例えば図 4.1 の例で、往信文「③ ですが、空調のフィルターにホコリがたまっていたのが見えてしまいました。」と返信文「③ 直ちに清掃担当者に申し伝え、清掃指示及び今後のチェック強化を徹底いたしました。」は共有する内容語を一つも持たないにも関わらず文対応が存在する。このように、文対応がある文同士は必ずしも類似の表現を用いているとは限らない。そのため、単純な文の類似度によらない推定手法が必要となる。
- (2) **文の出現順序と文対応の出現位置は必ずしも一致しない:** 例えば図 4.1 の例で対応が逆転している（文対応を示す直線が交差している）ように、返信文書の書き手は往信文書の並びと対応させて返信文書を書くとは限らない。そのため、文書中の出現位置に依存しない推定手法が必要となる。

本研究では、以上の特徴を踏まえて文対応の自動推定を実現するために、本課題を文対応の有無を判定する二値分類問題と考える。すなわち、存在しうる全ての文対応（例えば図 4.1 であれば $4 \times 5 = 20$ 通り）のそれぞれについて文対応が存在するかを判定する分類器を作成する。分類器の作成にあたって、文種類（対象がレビュー文書・応答文書であれば「ポジティブな感想」「ネガティブな感想」や「ほめへの感謝」「お詫び」など）を利用する手法を提案する。まず、最初にそれぞれの文書の文種類を推定し、文対応推定の素性として用いる手法を提案する。次に、この手法を発展させた文種類と文対応を同時に

推定するモデルを提案し、より高い性能で文対応の推定が実現できることを示す。

4.2 文種類・文対応

提案手法の説明に先立ち、最初に推定すべき文対応について改めて定義する。次に、文対応の推定に用いる素性の一つである文種類について説明する。最後に、実際のレビュー文書・応答文書集合における文種類・文対応に関わる諸データを示す。

4.2.1 文対応

文対応推定のための提案手法を説明する前に、本研究課題で扱う**文対応**について改めて定義する。本研究では文対応を次のように定義し、この関係の有無を推定することを目的とする。

定義3 (文対応). 文書とそれに対する返信文書が与えられた時、ある返信文が返信先文書のある文を原因として生起している場合、それらの文の間には**文対応**関係があるとする。

なお、文対応は返信文書中の1文から、往信文書中の複数の文へ対応することを許す。また、返信文書中の異なる文から、往信文書中の同一の文への対応も許す。

具体的に図4.1を例に説明する。例えば「④ ご朝食につきましてはお気に召していただけただようで大変光栄に存じます。」と言う返信文は、「② 朝食メニューが新しくなったのが良かったです。」と言う往信文を原因として書かれた文であるため、文対応を持つ。同様に、返信文「② しかしながらこの度のご宿泊の折には、空調フィルターのホコリによりご不快な思いをおかけいたしまして大変申し訳ございませんでした。」は往信文「③ ですが、空調のフィルターにホコリがたまっていたのが見えてしまいました。」「④ ホコリは一旦目に付くと気になってしまうので、どうかチェックして頂けたらと思います。」を原因として書かれた文であるため、両方の対とも文対応を持つ。

4.2.2 文種類

次に、文対応推定に用いる素性の一つである文種類についても説明する。本研究では、「ある文がどのような目的で書かれているかによる分類」を**文種類**と定義する。

どのような文種類の集合を用いるかは、対象とする文書と目的によって異なる。例えば、一般的なメールと応答であれば「挨拶」「質問」「謝罪」「回答」などの文種類を考えることができる。本研究では、対象をレビュー文書と応答文書としているため、これらの文書における文種類の分類を行っている大沢他(2010b)の先行研究を元に文種類を定義した。大沢他(2010b)は本研究の実験でも用いる楽天トラベルの「クチコミ・お客さまの声」を分析し、レビュー文を8種類、応答文を14種類に分類している。本研究では、目的である文対応の特定に特化するため、文種類が特定できれば文対応の特定が容易になるように、レビュー文を12種類、応答文を20種類に再分類した。この分類を表4.1及び表4.2に示す*1*2。また、大沢他(2010b)の分類からの主な変更点とその理由を以下に示す。

(1) **レビュー文種類 – <ポジ/ネガ感想> の追加**: 一文にポジティブ・ネガティブな感想を双方含むレ

*1 レビュー文書の中には、末尾に「【ご利用の宿泊プラン】」に続いて宿泊プランの名称が書かれている文が存在した。この記述は、おそらく楽天トラベルのレビューを投稿する際に自動で挿入される文であると考えている。

*2 なお、「その他」は例えば文書が英語で書かれているため分類が不可能であった文などである。

表 4.1: レビュー文書を構成する文の種類 (大沢他 (2010b) による分類を参考に再構成)

分類	レビュー文の例
<宿泊説明> 宿泊理由、経緯や部屋タイプ等の説明	4月2日(金)、3日(土)の1泊2日で家族5人で宿泊しました。
<感謝・応援> 宿泊施設への感謝・御礼・応援	<ul style="list-style-type: none"> ・大変お世話になりました。 ・頑張ってください。
<プラン名> 利用したプランの名前	【ご利用の宿泊プラン】 ☆2名様でお得！現金お支払プラン☆【シングルエキストラ】
<再泊・推薦意向> 再度の宿泊意向又は他者へ推薦する旨	<ul style="list-style-type: none"> ・またお伺いしたいと思います。 ・会社の同僚にも薦めておきました。
<不再泊・不薦意向> 再度宿泊しない又は推薦しない旨	次は、泊まらないかな～
<ポジティブ感想> 施設提供サービス等が良かった旨の記述	外観、ロビー部屋共に宿泊料金からすればとても満足できるものです。
<ポジ/ネガ感想> ポジ・ネガ感想双方を含む	大分駅から商店街を通過って向かうとホテルの看板が非常に小さくわかり難いですが、部屋は快適です。
<ネガティブ感想> 施設提供サービス等が悪かった旨の記述	またモーターのような音もうるさくよく眠れなかった。
<ニュートラル感想> 施設提供サービス等が普通だった旨の記述	アメニティも普通です。
<要求・要望> 施設への希望や要求	あとは、部屋の窓を開けられるようにして欲しいです。
<情報追加> 他ユーザへの追加的情報の提示	地下鉄が使えれば品川駅より都営浅草線高輪台駅がお勧め。
<個別事情の説明> 施設サービスとは直接関係ない個人の事情	お風呂を利用する時間がなかったのが実に残念です。

表 4.2: 応答文書を構成する文の種類 (大沢他 (2010b) による分類を参考に再構成)

分類	応答文の例
<宿泊御礼> 宿泊のお礼	この度は当施設をご利用いただきましてありがとうございます。
<投稿御礼> 投稿のお礼	ご感想もお寄せ下さり重ねて御礼申し上げます。
<結びでの御礼> 結びでの御礼	ありがとうございました。
<心掛け・決意> 今後の心掛けや決意の表明	今後も、お客様に喜んでいただけるよう、現状に甘んじる事なくサービス・施設の向上に努めていく所存です。
<再泊願ひ> 今後ご愛願頂きたい気持ちの表明	またのご利用を心よりお待ち申し上げます。
<署名・フッター> 返答者の署名・文書のフッター	・宿泊課 ** ・☆☆☆☆☆☆☆☆☆☆☆☆
<定型的挨拶> 頭語・結語・時候/結びの挨拶	2009 年は格別の御厚情を賜り、厚く御礼を申し上げます。
<ほめへの感謝> ほめの言葉への感謝	お褒めの言葉を頂戴いたしまして、大変嬉しく存じます。
<お詫び> 指摘に対する謝罪	心より、お詫び申し上げます。
<恥じ入り> 恥ずべき行為への反省	連休中ともあり、お待ち頂く可能性がある旨をお客様へ事前にご案内すべきだったと深く反省しております。
<提案・指摘感謝> 提案や指摘への感謝	ルームキーにつきましては、貴重なご意見有難う御座います。
<事実述べ> 要望・苦情を受けての事実や現況の明示	このプランは建物の構造上窓が遮られてしまいますが、他のお部屋と設備は変わりません。
<具体的対応明示> 具体的な改善・対応策の明示	パソコン本体との接触するパーツに不具合があったため、パーツを交換させていただきました。
<抽象的対応明示> 抽象的な改善・対応策の明示	今後も同じことがないように心がけてまいります。
<具体的検討明示> 指摘事項等に対する具体的な検討の明示	もう少し厚くできるか検討させていただきます。
<抽象的検討明示> 指摘事項等に対する抽象的な検討の明示	お客様からいただいたご意見を検討してまいります。
<対処明示> 要望・苦情に対して既存の対応方法を明示	数に限りがございますが、追加のハンガーでよろしかったらフロントにてお貸し出しがございます。
<了承願ひ> 現況では改善できないため、了承を請うもの	高層階に位置していないお部屋タイプもございますので、重ねてご理解の程お願い申し上げます。
<対話> レビュー内容を受けての対話	サトウキビの収穫体験はいかがでしたか？
<情報追加> 利用者・他ユーザへの追加的情報の提示	ご朝食には、パン・おにぎり・サラダ・フルーツ・ヨーグルト・スープなど、種類豊富な食材をご用意しております。

ビュー文は、複数の応答文と対応することがあるため。

- (2) **応答文種類 – <対応明示> <検討明示> の具体性による細分化:** 具体性のある対応・検討明示文はそれぞれレビュー文で書かれた一つの事情と対応するが、抽象的なものはレビュー文で書かれた複数の事情と対応することがあるため。

4.2.3 文種類と文対応データの分析

本節では、実際のレビュー文書と応答文書のデータについて、文種類と文対応に関連したデータを示す。データには「楽天データ公開^{*3}」に収録されている楽天トラベルのレビューを用いた。収録されているレビュー 348,564 件のうち、レビュー文書・応答文書の双方が存在する 276,562 件からランダムサンプリングした 1,000 文書対を用いた。この各文書を簡易的なヒューリスティック（詳しくは第 A 章で説明する）によって文単位に分割し、レビュー文 4,813 文・応答文 6,160 文を得た。

文種類データの分析

まず、文種類に関わるデータを示す。得られたレビュー文・応答文に対して、表 4.1, 4.2 に示す文種類の分類に基づき、人手で文種類をタグ付けした。それぞれの文種類の分布は表 4.3, 4.4 の通りであった。

次に、レビュー文書・応答文書ごとに、ある文種類 t_1 が与えられた下での文種類 t_2 の遷移確率 $P(t_2|t_1)$ について調査した。このうち、10 事例以上観察され、特に遷移確率が高い上位 10 例を表 4.5, 4.6 に示す。表 4.5, 4.6 より、文種類 t_2 の予測において、1 つ前の文種類 t_1 の情報が有用な場合があることが分かる。

文対応データの分析

次に、文対応に関わるデータを示す。得られたレビュー文書・応答文書対に対して、定義 3 に示す定義に基づいて文対応の有無をタグ付けした。その結果、1,000 文書対全体では 4,492 通りの文対応（1 文書対あたり 4.49 文対応）が得られた。

まず、文種類について、各文種類の出現数と各文種類ごとに文対応がどの程度存在するかを調査したデータを表 4.7, 4.8 に示す。表中の「対応（平均数）」は一文から見たときの平均対応文数を示しており、「対応（存在率）」は一つでも対応が存在する割合を示している。表 4.7, 4.8 より、例えばレビュー文種類では <ネガティブ感想> や <要求・要望> が、応答文種類では <お詫び> や <具体的対応明示> などの文種類で対応存在率が高いなど、文種類によって対応の平均数や対応存在率が大きく異なることが分かる。第 1 章でサービスの失敗に対する顧客対応の重要性について述べたが、特に <ネガティブ感想>、<要求・要望> に対する文対応の存在率はそれぞれ 83.70%, 79.62% であることから、楽天トラベルの応答文書においても顧客対応が重要視されていることを示唆している。また、文対応が存在する文種類の組み合わせの上位 15 例を表 4.9 に示す。表 4.9 より、これらの文対応については、文種類が特定できれば文対応推定に有用であると言える。

次に、文対応の遷移確率を示す。まず、応答文を固定した際に、同じ応答文に対するレビュー文間での遷移確率を表 4.10 に示す。同様に、レビュー文を固定した際に、同じレビュー文に対する応答文間での遷移確率を表 4.11 に示す。表 4.10, 4.11 より、いずれの場合とも「なし」から「なし」及び「あり」から「あり」への遷移が比較的多いことから、同じレビュー文（応答文）に対しては位置的に連続する

^{*3} 楽天データ公開: <http://rit.rakuten.co.jp/rdr/>。データは 2010 年時点の公開データを用いた。

表 4.3: レビュー文での文種類の分布

文種類	出現数
宿泊説明	386
感謝・応援	121
プラン名	196
再泊・推薦意向	382
不再泊・不薦意向	16
ポジティブ感想	2,263
ポジ/ネガ感想	86
ネガティブ感想	816
ニュートラル感想	72
要求	260
情報追加	71
個別事情の説明	129
その他	15
合計	4,813

表 4.4: 応答文での文種類の分布

文種類	出現数
宿泊御礼	1,050
投稿御礼	324
結びでの感謝	88
心掛け・決意	407
再泊願い	1,187
署名・フッター	625
定型的挨拶	32
ほめへの感謝	623
お詫び	365
恥じ入り	10
提案・指摘感謝	80
事実述べ	128
具体的対応明示	131
抽象的対応明示	86
具体的検討明示	26
抽象的検討明示	86
対処明示	102
了承願い	20
対話	152
情報追加	613
その他	25
合計	6,160

表 4.5: レビュー文種類の遷移確率例（遷移確率の高い上位 10 例）

レビュー文種類		遷移確率
文種類 t_1	→ 文種類 t_2	$P(t_2 t_1)$
宿泊説明	→ ポジティブ感想	70.57%
ネガティブ感想	→ ネガティブ感想	61.53%
ポジティブ感想	→ ポジティブ感想	60.49%
再泊・推薦意向	→ プラン名	40.00%
個別事情の説明	→ ポジティブ感想	37.84%
ポジ/ネガ感想	→ ポジティブ感想	35.00%
ニュートラル感想	→ ポジティブ感想	31.25%
情報追加	→ ポジティブ感想	30.51%
感謝・応援	→ プラン名	28.81%
ポジ/ネガ感想	→ ネガティブ感想	28.33%

表 4.6: 応答文種類の遷移確率例 (遷移確率の高い上位 10 例)

応答文種類		遷移確率
文種類 t_1	→ 文種類 t_2	$P(t_2 t_1)$
署名・フッター	→ 署名・フッター	98.32%
結びでの感謝	→ 署名・フッター	91.18%
心掛け・決意	→ 再泊願ひ	66.42%
抽象的検討明示	→ 再泊願ひ	44.58%
再泊願ひ	→ 署名・フッター	44.35%
具体的検討明示	→ 再泊願ひ	44.00%
情報追加	→ 情報追加	43.62%
対話	→ 対話	39.47%
提案・指摘感謝	→ 再泊願ひ	38.96%
抽象的対応明示	→ 再泊願ひ	38.37%

表 4.7: レビュー文ごとの対応数・対応存在率

レビュー文種類	文数	対応 (平均数)	対応 (存在率)
宿泊説明	386	0.33	25.65%
感謝・応援	121	0.09	9.09%
プラン名	196	0.01	0.51%
再泊・推薦意向	382	0.20	14.40%
不再泊・不薦意向	16	0.13	12.50%
ポジティブ感想	2263	0.88	62.48%
ポジ/ネガ感想	86	1.78	76.74%
ニュートラル感想	72	0.26	20.83%
ネガティブ感想	816	1.97	83.70%
要求・要望	260	1.56	79.62%
情報追加	71	0.34	22.54%
個別事情の説明	129	0.57	37.21%
その他	15	0.07	6.67%
レビュー文全体	4813	0.93	54.00%

表 4.8: 応答文ごとの対応数・対応存在率

応答文種類	文数	対応 (平均数)	対応 (存在率)
宿泊御礼	1050	0.10	9.24%
投稿御礼	324	0.04	1.54%
結びでの感謝	88	0.00	0.00%
心掛け・決意	407	0.30	17.69%
再泊願ひ	1187	0.03	2.27%
署名・フッター	625	0.00	0.00%
定型的挨拶	32	0.09	6.25%
ほめへの感謝	623	2.26	97.59%
お詫び	365	2.27	96.99%
恥じ入り	10	2.00	100.00%
提案・指摘感謝	80	1.70	90.00%
事実述べ	128	1.83	99.22%
具体的対応明示	131	1.83	100.00%
抽象的対応明示	86	2.86	98.84%
具体的検討明示	26	1.65	100.00%
抽象的検討明示	86	1.97	97.67%
対処明示	102	1.47	99.02%
了承願ひ	20	1.55	100.00%
対話	152	0.61	51.32%
情報追加	613	0.99	80.10%
その他	25	0.60	36.00%
応答文全体	6160	0.73	38.94%

表 4.9: 文対応が存在する文種類の組み合わせ (対応率上位 15 例)

両文種類		対応率	数
要求・要望	- 提案・指摘感謝	86.76%	59
ポジ/ネガ感想	- 対処明示	84.62%	11
要求・要望	- 抽象的検討明示	84.34%	33
ネガティブ感想	- 了承願い	83.33%	25
要求・要望	- 抽象的対応明示	75.00%	33
要求・要望	- 具体的検討明示	73.68%	70
ポジ/ネガ感想	- 具体的対応明示	73.68%	12
ネガティブ感想	- お詫び	70.66%	667
ネガティブ感想	- 恥じ入り	69.23%	18
ネガティブ感想	- 抽象的対応明示	68.28%	198
ポジ/ネガ感想	- お詫び	67.14%	47
ポジ/ネガ感想	- 抽象的対応明示	66.67%	14
ポジティブ感想	- ほめへの感謝	66.55%	1,317
要求・要望	- 事実述べ	64.52%	40
ポジ/ネガ感想	- ほめへの感謝	63.64%	21

表 4.10: 文対応の遷移確率 (レビュー文基準)

文対応の有無		遷移確率
a	\rightarrow b	$P(b a)$
なし	\rightarrow なし	91.92%
なし	\rightarrow あり	8.08%
あり	\rightarrow なし	49.10%
あり	\rightarrow あり	50.90%

表 4.11: 文対応の遷移確率 (応答文基準)

文対応の有無		遷移確率
a	\rightarrow b	$P(b a)$
なし	\rightarrow なし	88.12%
なし	\rightarrow あり	11.88%
あり	\rightarrow なし	60.43%
あり	\rightarrow あり	39.57%

応答文 (レビュー文) が文対応を持つことが多いことを示唆している。

更に、文対応が交差する割合を示す。交差割合の計算は、各文書対において「文書対内において交差を持つ文対応の数 / 文書対内における全ての文対応の数」により求めた。結果、交差割合の平均は 0.249 であることから、本データにおいても文の出現順序と文対応の出現位置は必ずしも一致しないことが確認できた。

最後に、文対応の有無別に文の類似度を比較した結果を示す。類似度の計算にあたり、まず文をベクトル空間モデル (Vector Space Model) によりベクトル化した。ベクトル空間モデルは、データセット中の全単語種類数を V とした時、文や文書を各次元が単語と対応するような V 次元ベクトル空間で表現するモデルであり、例えば文書検索などで文書のベクトル化が必要な際に広く用いられている (Manning & Schütze, 1999)。ここで、ベクトルの各次元の値は、次元に対応する単語が文や文書において出現した頻度となる。また、ベクトル化した文ベクトル $\mathbf{v}_1, \mathbf{v}_2$ の類似度はコサイン類似度 $\mathbf{v}_1 \cdot \mathbf{v}_2 / \|\mathbf{v}_1\| \|\mathbf{v}_2\|$ により計算した。文対応の有無別にコサイン類似度の分布を表したヒストグラムを図 4.2, 4.3 に示す (なお、コサイン類似度が 0.3 - 1.0 である文対応の割合は少なかったため省略している)。なお、コサイン類似

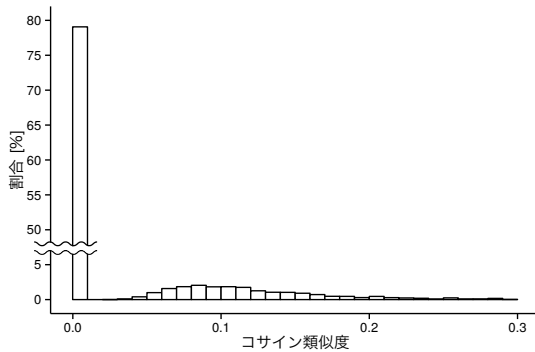


図 4.2: コサイン類似度の分布 (文対応なし)

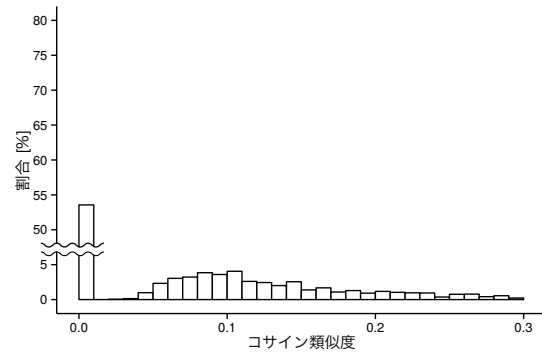


図 4.3: コサイン類似度の分布 (文対応あり)

度は、各文における stop-word*4 を除く単語の出現頻度を値を持つベクトルを用いて計算した値である。文対応を持つ文間の方が比較的高いコサイン類似度が高い傾向がある一方、文対応が存在する文対のうち 53.56% はコサイン類似度が 0 であった。そのため、本データにおいても対応する文同士は必ずしも類似しないことが確認できた。

4.3 基盤技術

本研究課題での提案手法では CRF と呼ばれる確率モデルを用いて文対応関係の推定を行う。そこで、本節では CRF に必要な諸概念の導入と CRF 一般の説明を行い、続いて CRF のバリエーションである Linear-chain CRF、2D CRF について説明する。

4.3.1 Conditional Random Fields

Conditional Random Fields (CRF; 条件付確率場) は、Lafferty et al. (2001) が提案した教師あり確率的識別モデルの一つである。CRF をはじめとする確率的識別モデルでは、次の式 (4.1) で表すように入力系列 \mathbf{x} の条件のもとでの確率が最大となる出力系列 $\hat{\mathbf{y}}$ を出力として採用する。

$$\hat{\mathbf{y}} = \underset{\mathbf{y}}{\operatorname{argmax}} P(\mathbf{y}|\mathbf{x}) \quad (4.1)$$

以下、まず最初に CRF の説明に必要な素性の概念について説明する。次に、CRF のモデル化に必要な対数線形モデルについて説明する。続いて、具体的にどのように $P(\mathbf{y}|\mathbf{x})$ を因数分解するかを説明するため、因子グラフの概念を導入する。最後に、 $P(\mathbf{y}|\mathbf{x})$ の因数分解の方法によるバリエーションである Linear-chain CRF、2D CRF を説明する。

素性

CRF をはじめとする識別モデルでは、文や単語と言った入力を様々な側面から数値として表現するために素性 (feature) という単位を用いる。ここで CRF の説明に必要な素性の概念について簡単に説明する。

*4 本研究では、品詞が「助詞」「助動詞」「記号」の単語を stop-word とした。

例えば、2つの文が与えられた時にそれらの文が対応しているかを推定する単純な識別器を考える。ここでは入力 \mathbf{x} が文の対、出力 \mathbf{y} が文対応に当たる。このような文の対 \mathbf{x} を見ると、例えばそれぞれの文が持つ各単語と言う特徴、2つの文がどの程度類似しているかと言った特徴が存在する。このような特徴のとらえ方を素性と呼び、 \mathbf{x} をそれぞれの素性に合わせて数値化する関数を素性関数と呼ぶ。

対数線形モデル

CRF は対数線形モデルによってモデル化した $P(\mathbf{y}|\mathbf{x})$ を識別に用いる。そこで、対数線形モデルの導出に詳しい Ratnaparkhi (1997) の解説を参考に導出方法を説明する。

教師あり学習において、学習データが与えられた際にモデルを最適化する方法の中で最も単純な方法は最尤推定による方法である。最尤推定では学習データの各入力から各出力が最も出現しやすくなるようにモデルのパラメータを調整するため、確率モデルにおいては学習データ上での確率分布がモデル上の分布と一致する。しかし、最尤推定では学習データに出現しなかった入力を与えると確率値が0になってしまうと言う過学習が問題となる。特に自然言語処理で扱う課題では無数に存在するデータのごく一部しか用意できないため、この問題は顕著である。

この問題に対処するため、対数線形モデルでは「学習データで不明なところはできるだけ偏りが無いようにする」と言うエントロピー最大化を目的とした最尤法によってパラメータを決定する。今、可能な全入出力変数の集合をそれぞれ \mathcal{X}, \mathcal{Y} と置くと、エントロピーを求める式は

$$H(P) = - \sum_{\mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}} P(\mathbf{x}, \mathbf{y}) \log P(\mathbf{x}, \mathbf{y}) \quad (4.2)$$

である。ここで、対数線形モデルでは同時確率 $P(\mathbf{x}, \mathbf{y})$ を次の式 (4.3) のように分解する。なお、 $P'(\mathbf{x})$ は適当な確率分布である (後にモデルのパラメータとして吸収されるので特に限定する必要はない)。

$$P(\mathbf{x}, \mathbf{y}) = P'(\mathbf{x})P(\mathbf{y}|\mathbf{x}) \quad (4.3)$$

この式を用いると、 $H(P)$ の最大化は

$$\begin{aligned} \operatorname{argmax}_P H(P) &= \operatorname{argmax}_P - \sum_{\mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}} P(\mathbf{x}, \mathbf{y}) \log P(\mathbf{x}, \mathbf{y}) \\ &= \operatorname{argmax}_P - \sum_{\mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}} P'(\mathbf{x})P(\mathbf{y}|\mathbf{x}) \log(P'(\mathbf{x})P(\mathbf{y}|\mathbf{x})) \\ &= \operatorname{argmax}_P - \sum_{\mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}} P'(\mathbf{x})P(\mathbf{y}|\mathbf{x}) \log P(\mathbf{y}|\mathbf{x}) \end{aligned} \quad (4.4)$$

とできる。この式 (4.4) の $-\sum_{\mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}} P'(\mathbf{x})P(\mathbf{y}|\mathbf{x}) \log P(\mathbf{y}|\mathbf{x})$ を $H'(P)$ とする。

次に、最大化に先立ち、モデルの性質に沿った二つの制約を導入する。一つ目の制約は学習データに最適化するための制約であり、真の分布 P と学習データ上の分布 \tilde{P} における各素性関数の期待値が同一であると言う制約である。この制約は、各素性関数を f_i 、全素性関数の集合を \mathcal{F} と置くと、次の式 (4.5) で表現できる。

$$\sum_{\mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}} P(\mathbf{y}|\mathbf{x})f_i(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}} \tilde{P}(\mathbf{y}|\mathbf{x})f_i(\mathbf{x}, \mathbf{y}) \quad \forall f_i \in \mathcal{F} \quad (4.5)$$

二つ目の制約は、 P は確率関数であるから総和を取ると1になるという制約である。すなわち

$$\sum_{\mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}} P(\mathbf{x}, \mathbf{y}) = 1 \quad (4.6)$$

である。

以上をまとめると、対数線形モデルはエントロピーの最大化 (4.4) を制約 (4.5), (4.6) の元で実現するモデルであると言える。これによって最適パラメータの決定が頑健性を保ったまま実現する。

この最大化問題はラグランジュの未定乗数法によって解析的に解くことが可能である。ラグランジュ乗数を λ と置くと、この場合のラグランジュ関数 \mathcal{L} は次の式 (4.7) で表せる。

$$\mathcal{L} = H'(P) + \sum_{f_i \in \mathcal{F}} \lambda_i \left(\sum_{\mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}} P(\mathbf{y}|\mathbf{x}) f_i(\mathbf{x}, \mathbf{y}) - \sum_{\mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}} \tilde{P}(\mathbf{y}|\mathbf{x}) f_i(\mathbf{x}, \mathbf{y}) \right) + \lambda_0 \left(\sum_{\mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}} P(\mathbf{x}, \mathbf{y}) - 1 \right) \quad (4.7)$$

$H'(P)$ が最大となる点では、

$$\frac{\partial}{\partial P(\mathbf{y}|\mathbf{x})} \mathcal{L} = 0 \quad \forall \mathbf{x} \in \mathcal{X}, \forall \mathbf{y} \in \mathcal{Y} \quad (4.8)$$

が成り立つので、これを求めると

$$\frac{\partial}{\partial P(\mathbf{y}|\mathbf{x})} \mathcal{L} = -P'(\mathbf{x}) \log P(\mathbf{y}|\mathbf{x}) + \sum_{f_i \in \mathcal{F}} \lambda_i f_i(\mathbf{x}, \mathbf{y}) + \lambda_0 P'(\mathbf{x}) \quad (4.9)$$

である。0 と置いて整理すると

$$\begin{aligned} & -P'(\mathbf{x}) \log P(\mathbf{y}|\mathbf{x}) + \sum_{f_i \in \mathcal{F}} \lambda_i f_i(\mathbf{x}, \mathbf{y}) + \lambda_0 P'(\mathbf{x}) = 0 \\ \Leftrightarrow & \log P(\mathbf{y}|\mathbf{x}) = \frac{1}{P'(\mathbf{x})} \cdot \sum_{f_i \in \mathcal{F}} \lambda_i f_i(\mathbf{x}, \mathbf{y}) + \lambda_0 \\ \Leftrightarrow & P(\mathbf{y}|\mathbf{x}) = \exp \left\{ \frac{1}{P'(\mathbf{x})} \cdot \sum_{f_i \in \mathcal{F}} \lambda_i f_i(\mathbf{x}, \mathbf{y}) + \lambda_0 \right\} \\ \Leftrightarrow & P(\mathbf{y}|\mathbf{x}) = \frac{1}{Z_{\mathbf{x}}} \exp \left\{ \sum_{f_i \in \mathcal{F}} \theta_i f_i(\mathbf{x}, \mathbf{y}) \right\} \quad \left(\frac{1}{Z_{\mathbf{x}}} = \exp \lambda_0, \theta_i = \frac{\lambda_i}{P'(\mathbf{x})} \right) \end{aligned} \quad (4.10)$$

となり、対数線形モデルにおける $P(\mathbf{y}|\mathbf{x})$ を得ることができた。 $\frac{1}{Z_{\mathbf{x}}}$ は (4.6) を満足するための正規化項 (すなわち $Z_{\mathbf{x}} = \sum_{\mathbf{y}' \in \mathcal{Y}} \exp \sum_{f_i \in \mathcal{F}} \theta_i f_i(\mathbf{x}, \mathbf{y}')$)、 θ_i は素性 f_i の重みを表すパラメータである。対数線形モデルではこの式 (4.10) を用いて、入力 \mathbf{x} に対する出力 \mathbf{y} の確率を計算する。

モデルの式の導入が完了したので、続いて対数線形モデルにおいて学習時にパラメータを推定する方法を説明する。対数線形モデルでは式 (4.10) に登場する素性の重み θ_i がパラメータである。このパラメータの決定は一般に事後確率最大化 (MAP; Maximum A Posterior) によって行うが、まずは最尤推定の方法から説明し、続いて事後確率最大化による方法を説明する。

まず、学習データにおける入出力の組を $\langle x^{(j)}, y^{(j)} \rangle$ とすると、最尤推定では尤もらしい θ である $\hat{\theta}$ を次のようにして求める。

$$\hat{\theta} = \operatorname{argmax}_{\theta} \prod_j P(\mathbf{y}^{(j)}|\mathbf{x}^{(j)}; \theta) \quad (4.11)$$

$P(\mathbf{y}^{(j)}|\mathbf{x}^{(j)}; \theta)$ を対数化した対数尤度関数 \mathcal{L}_{θ} を求め、式 (4.10) の $P(\mathbf{y}|\mathbf{x})$ を代入すると

$$\begin{aligned}
\mathcal{L}_\theta &= \sum_j \log P(\mathbf{y}^{(j)}|\mathbf{x}^{(j)}; \theta) \\
&= \sum_j \log \frac{\exp\left(\sum_{f_i \in \mathcal{F}} \theta_i f_i(\mathbf{x}^{(j)}, \mathbf{y}^{(j)})\right)}{\sum_{\mathbf{y}' \in \mathcal{Y}} \exp\left(\sum_{f_i \in \mathcal{F}} \theta_i f_i(\mathbf{x}^{(j)}, \mathbf{y}')\right)} \\
&= \sum_j \left\{ \sum_{f_i \in \mathcal{F}} \theta_i f_i(\mathbf{x}^{(j)}, \mathbf{y}^{(j)}) - \log \sum_{\mathbf{y}' \in \mathcal{Y}} \exp\left(\sum_{f_i \in \mathcal{F}} \theta_i f_i(\mathbf{x}^{(j)}, \mathbf{y}')\right) \right\} \tag{4.12}
\end{aligned}$$

である。 \mathcal{L}_θ の凸性から \mathcal{L}_θ の最適点では、

$$\frac{\partial}{\partial \theta_k} \mathcal{L}_\theta = \sum_j \left\{ f_k(\mathbf{x}^{(j)}, \mathbf{y}^{(j)}) - \frac{\sum_{\mathbf{y}' \in \mathcal{Y}} f_k(\mathbf{x}^{(j)}, \mathbf{y}') \exp\left(\sum_{f_i \in \mathcal{F}} \theta_i f_i(\mathbf{x}^{(j)}, \mathbf{y}')\right)}{\sum_{\mathbf{y}' \in \mathcal{Y}} \exp\left(\sum_{f_i \in \mathcal{F}} \theta_i f_i(\mathbf{x}^{(j)}, \mathbf{y}')\right)} \right\} = 0 \tag{4.13}$$

が成立する。ここで、中括弧内の第2項目の分母 $\sum_{\mathbf{y}' \in \mathcal{Y}} \exp\left(\sum_{f_i \in \mathcal{F}} \theta_i f_i(\mathbf{x}^{(j)}, \mathbf{y}')\right)$ が $P(\mathbf{y}'|\mathbf{x}^{(j)})$ の正規化項 $Z_{\mathbf{x}^{(j)}}$ であることに注目すると、

$$\begin{aligned}
\frac{\partial}{\partial \theta_k} \mathcal{L}_\theta &= \sum_j \left\{ f_k(\mathbf{x}^{(j)}, \mathbf{y}^{(j)}) - \sum_{\mathbf{y}' \in \mathcal{Y}} f_k(\mathbf{x}^{(j)}, \mathbf{y}') \frac{1}{Z_{\mathbf{x}^{(j)}}} \exp\left(\sum_{f_i \in \mathcal{F}} \theta_i f_i(\mathbf{x}^{(j)}, \mathbf{y}')\right) \right\} \\
&= \sum_j \left\{ f_k(\mathbf{x}^{(j)}, \mathbf{y}^{(j)}) - \sum_{\mathbf{y}' \in \mathcal{Y}} P(\mathbf{y}'|\mathbf{x}^{(j)}) f_k(\mathbf{x}^{(j)}, \mathbf{y}') \right\} \quad (\text{式 (4.10) を利用}) \\
&= \sum_j f_k(\mathbf{x}^{(j)}, \mathbf{y}^{(j)}) - \sum_j \sum_{\mathbf{y}' \in \mathcal{Y}} P(\mathbf{y}'|\mathbf{x}^{(j)}) f_k(\mathbf{x}^{(j)}, \mathbf{y}') \tag{4.14}
\end{aligned}$$

と変形できる。式 (4.14) の第1項目は素性 f_k の学習データにおける総和（素性関数が二値関数であれば、当該素性に該当する素性の数と一致する）、第2項目はモデル P における素性 f_k の期待値と解釈できる。最終的に、各 θ_k について式 (4.14) を 0 と置き、 θ について解けば最適であるパラメータ $\hat{\theta}$ を求めることができる。この最適パラメータ $\hat{\theta}$ の存在性及び唯一性は Ratnaparkhi (1997) によって示されている。

この期待値の計算は、forward-backward アルゴリズムや信念伝播法を適用することによって効率的に計算できる。この計算方法により、可能な \mathbf{y}' の組み合わせの全列挙による計算時間の肥大化を避けられる利点がある。

ただし、実際には最尤推定でパラメータを最適化してしまうと、パラメータに極端な値が設定されるという過学習の問題が生じる可能性がある。そこで、パラメータ θ に事前分布 $P(\theta)$ を仮定した事後確率最大化によるパラメータの正則化を行う方法について説明する。この場合、式 (4.11) は

$$\hat{\theta}_{\text{MAP}} = \underset{\theta}{\operatorname{argmax}} \prod_j P(\mathbf{y}^{(j)}|\mathbf{x}^{(j)}; \theta) P(\theta) \tag{4.15}$$

となる。事前分布に正規分布を仮定すると、対数尤度関数 \mathcal{L}'_θ は次の式 (4.16) のようになる。

$$\mathcal{L}'_\theta = C \cdot \sum_j \log P(\mathbf{y}^{(j)}|\mathbf{x}^{(j)}; \theta) - \frac{1}{2} \sum_k \theta_k^2 \tag{4.16}$$

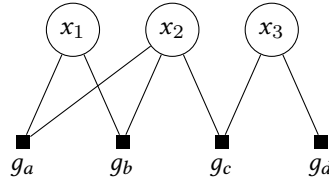


図 4.4: $G(x_1, x_2, x_3) = g_a(x_1, x_2)g_b(x_1, x_2)g_c(x_2, x_3)g_d(x_3)$ と対応する因子グラフの例

この式における第 2 項目 $-\frac{1}{2} \sum_k \theta_k^2$ が、大きすぎる θ_k に対するペナルティを与えることにより正則化が実現する。なお、 C は事前分布 $P(\theta)$ を決定する数値であり、正則化に伴うハイパラメータと言える。これも式 (4.14) と同様に θ_k で微分すると

$$\frac{\partial}{\partial \theta_k} \mathcal{L}'_{\theta} = C \left\{ \sum_j f_k(\mathbf{x}^{(j)}, \mathbf{y}^{(j)}) - \sum_j \sum_{\mathbf{y}' \in \mathcal{Y}} P(\mathbf{y}' | \mathbf{x}^{(j)}) f_k(\mathbf{x}^{(j)}, \mathbf{y}') \right\} - \theta_k \quad (4.17)$$

である。

式 (4.14) 及び (4.17) における最適パラメータ $\hat{\theta}$ は、L-BFGS 法 (Byrd et al., 1994) などの準ニュートン法を用いて計算することができる。

4.3.2 因子グラフ

問題によっては観測変数 \mathbf{x} と隠れ変数 \mathbf{y} は複数の変数からなる。この場合、これらの全変数間の依存関係を考慮した $P(\mathbf{y}|\mathbf{x})$ を直接学習しようとするるとスパースネスの問題が発生する。そこで、変数間に適度な独立性を仮定することで、スパースネスの問題を回避する。変数間の独立性を表現するために、以下では因子グラフ (factor graph) (Kschischang et al., 2001) を用いる。そのため、本節で因子グラフについて説明する。

因子グラフは、Kschischang et al. (2001) が提案した、次の式 (4.18) に示す g_x のように複数変数 x_1, \dots, x_n に依存する関数をどのように因数分解するかを示すための二部グラフである。ここで、 g_j は因数分解後の局所関数、 X_j は関数 g_j の引数となる $\{x_1, \dots, x_n\}$ の部分集合を示す。

$$G(x_1, \dots, x_n) = \prod_j g_j(X_j) \quad (4.18)$$

具体的な因子グラフの定義は以下の通りである。

定義 4 (因子グラフ (Kschischang et al., 2001)). 因子グラフとは、式 (4.18) における因子分解の構造を表現した二部グラフである。二部グラフは、変数ノード x_j 、因子ノード g_j から構成される。変数ノード x_i と因子ノード g_j 間の枝は、 x_i が g_j の引数であるとき、かつそのときに限り存在する。

実際の因子グラフの例を図 4.4 に示す。因子グラフでは、変数を円形ノード、因数分解で分解した各因子を正方形ノードで表す。具体的に、図 4.4 は、次の式 (4.19) に示す $G(x_1, x_2, x_3)$ の因数分解と対応している。

$$G(x_1, x_2, x_3) = g_a(x_1, x_2)g_b(x_1, x_2)g_c(x_2, x_3)g_d(x_3) \quad (4.19)$$

また、因子グラフを用い、観測変数と隠れ変数の双方を含む確率関数の因数分解を表現することもできる。この場合、一般的には観測変数を有色の円形ノード (本論文では灰色のノード)、隠れ変数を無色

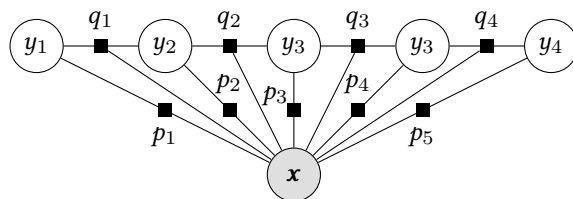


図 4.5: Linear-chain CRF の因子グラフ表現

の円形ノードで表す。実際の例を図 4.5 に示す。図 4.5 は、観測変数 \mathbf{x} 、隠れ変数 \mathbf{y} を引数に持つような確率関数 $P(\mathbf{y}|\mathbf{x})$ を次のように因数分解することを示している。

$$P(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{|\mathbf{y}|} p_i(\mathbf{x}, y_i) \cdot \prod_{j=1}^{|\mathbf{y}|-1} q_j(\mathbf{x}, y_j, y_{j+1}) \quad (4.20)$$

以降、因子グラフを数式で表現するために、因子グラフ \mathcal{G} を、変数ノードの集合 \mathcal{V} ・因子ノードの集合 \mathcal{F} ・エッジの集合 \mathcal{E} の 3 つを用いて $\mathcal{G} = (\mathcal{V}, \mathcal{F}, \mathcal{E})$ のように表現する。なお、集合の表記を簡便にするため次のような記法を導入する。例えば、 $\{y_i\}_{i=1}^5$ のように書いた場合、これは $\{y_i \mid 1 \leq i \leq 5 \wedge i \in \mathbb{N}\} = \{y_1, \dots, y_5\}$ を意味する。この記法を用いれば、図 4.5 及び式 (4.20) の因子グラフを $\mathcal{G} = (\mathcal{V}, \mathcal{F}, \mathcal{E})$ とすると $\mathcal{V}, \mathcal{F}, \mathcal{E}$ は次のように表現できる。

$$\mathcal{V} = \{\mathbf{x}\} \cup \{y_i\}_{i=1}^5 \quad (4.21)$$

$$\mathcal{F} = \{p_i\}_{i=1}^5 \cup \{q_j\}_{j=1}^4 \quad (4.22)$$

$$\mathcal{E} = \{(p_i, \mathbf{x})\}_{i=1}^5 \cup \{(p_i, y_i)\}_{i=1}^5 \cup \{(q_j, \mathbf{x})\}_{j=1}^4 \cup \{(q_j, y_j)\}_{j=1}^4 \cup \{(q_j, y_{j+1})\}_{j=1}^4 \quad (4.23)$$

4.3.3 Linear-chain Conditional Random Fields (Linear-chain CRF)

以上に説明した対数線形モデルと因子グラフの概念を用いると、CRF は次のように説明できる。すなわち、CRF は $P(\mathbf{y}|\mathbf{x})$ に対して適度な独立性を仮定し、 $P(\mathbf{y}|\mathbf{x})$ を対数線形モデルによりモデル化した識別モデルである。CRF には $P(\mathbf{y}|\mathbf{x})$ の独立性をどのように仮定するかによっていくつかのバリエーションが存在する。まず最初に、CRF のバリエーションとなるモデルとして、CRF の最も広く使われている形である Linear-chain Conditional Random Fields (Linear-chain CRF) について説明する*5。

Linear-chain CRF では、隠れ変数と観測変数がそれぞれ 1 次元の系列である場合を対象とする。具体的には、先に挙げた図 4.5 で示す因子グラフ上で、各因子関数に対数線形モデルを採用したものである。この場合、グラフは木構造であるため信念伝播法によって効率良く確率推定が可能である。以下、本研究では必要に応じて Linear-chain CRF を L-CRF と略す。図 4.5 を見ると、Linear-chain CRF では、因子関数は「隠れ変数と観測変数」を結ぶ因子 p_i と、「ある位置の隠れ変数と次の位置の隠れ変数」を結ぶ因子 q_j の 2 種類に分けられることが分かる。これはそれぞれ「隠れ変数と観測変数の値を考慮」「隠れ変数間の遷移を考慮」と言う方略に基づいた因数分解と言える。

*5 なお、ここで解説する Linear-chain CRF は遷移の考慮を 1 つ前だけに限定しているが、これは First order Linear-chain CRF と言う Linear-chain CRF の特別な場合である。本研究では Linear-chain CRF と言った場合は First order Linear-chain CRF に限ることとするが、例えば 2 つ前も含める Second order Linear-chain CRF などのバリエーションが存在する。

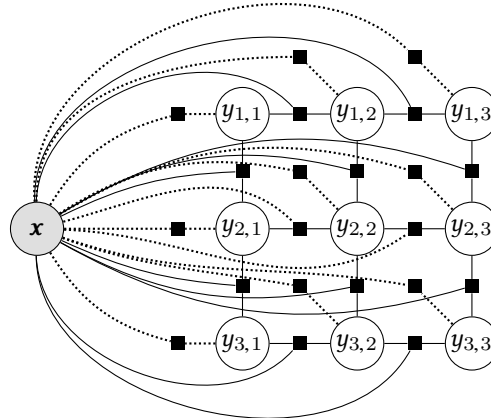


図 4.6: 2D CRF の因子グラフ表現

4.3.4 2D Conditional Random Fields (2D CRF)

次に、CRF の異なるバリエーションである 2D Conditional Random Fields (2D CRF) について説明する。2D CRF は、Zhu et al. (2005) が提案した二次元系列の隠れ変数を CRF によって一度に推定する識別モデルである*6。Linear-chain CRF と同様、各因子関数に対数線形モデルの $P(\mathbf{y}|\mathbf{x})$ を用いる。

2D CRF の構造を表す因子グラフを図 4.6 に示す。2D CRF では、「ある位置の隠れ変数と上下左右に隣接する隠れ変数」を結ぶ因子と (図 4.6 において、実線で接続される因子)、「隠れ変数と各観測変数」を結ぶ因子 (図 4.6 において、点線で接続される因子) の 2 種類が存在する。これらは「上下左右に隣接する隠れ変数間の値を考慮」「隠れ変数と各観測変数の値を考慮」と言う方略に基づいた因数分解と言える。Linear-chain CRF との大きな違いは、Linear-chain CRF では隠れ変数系列を一次的に結合していたのに対し、2D CRF では隠れ変数系列を二次元的に結合している点である。なお、図中では観測変数 \mathbf{x} を一つにまとめているが、 \mathbf{x} を分解し、それぞれを特定の因子ノードに接続しても良い。

4.4 提案手法

本節では、4.2.3 節で観察した文種類・文対応の特徴を踏まえ、4.3 節で導入した基盤技術に基づいた提案手法の説明を行う。まず、文対応推定問題の定式化を行う。次に、Linear-chain CRF による文種類推定手法を示し、続いて Linear-chain CRF, 2D CRF による文対応推定手法を順次示す。最後に、Linear-chain CRF による文種類推定モデル・2D CRF による文対応推定モデルを統合したモデルについて説明する。

*6 なお、本研究では隠れ変数が二次元系列である場合のモデルを 2D CRF と呼んでいるが、観測変数が二次元系列のモデルを 2D CRF と呼ぶ文献も存在する (Quattoni et al., 2004)。例えば Quattoni et al. (2004) は、観測変数が二次元系列の CRF を画像認識に用いている (この場合、観測変数は画像であるため二次元、隠れ変数は推定された物体名であるため一つ)。本研究では、隠れ変数が二次元系列である場合のモデルのみを扱う。

4.4.1 問題の定式化

提案手法の説明に先立ち、問題の定式化を行う。本研究の対象とする文書対は顧客からのメールやレビュー文書と販売者からの応答文書の対であるが、以下に示す手法は対をなす二文書であればこれらの文書対に限らず適用することができる。そのため、以下では顧客によるメール・レビュー文書、販売者による応答文書をより一般化して往信文書、返信文書と呼ぶ。

説明のために、以下の記号を導入する。ここで、文対応 $y_{i,j}$ は往信文 x^o_i から返信文 x^r_j の間に文対応関係が存在すれば 1、しなければ 0 の二値を取る変数である。

- N 往信文の文数
- M 返信文の文数
- \mathbf{x}^{org} 往信文の列 $x^o_1, x^o_2, \dots, x^o_N$
- \mathbf{x}^{rep} 返信文の列 $x^r_1, x^r_2, \dots, x^r_M$
- \mathbf{t}^{org} 各往信文の種類列 $t^o_1, t^o_2, \dots, t^o_N$
- \mathbf{t}^{rep} 各返信文の種類列 $t^r_1, t^r_2, \dots, t^r_M$
- $y_{i,j}$ 往信文 x^o_i と返信文 x^r_j の間における文対応存在の有無 (二値)
- \mathbf{y} 全ての文対応 $y_{i,j}$ からなる集合 $\{y_{i,j} \mid 1 \leq i \leq N, 1 \leq j \leq M\}$

各文間の文対応を推定する問題は、二文書 $\mathbf{x}^{\text{org}}, \mathbf{x}^{\text{rep}}$ を観測変数、文対応 \mathbf{y} を隠れ変数と考え、 $\mathbf{x}^{\text{org}}, \mathbf{x}^{\text{rep}}$ から \mathbf{y} を推定する問題と考えることができる。なお、定式化にあたって、文対応は多対多関係を持つものとし、一対一などの制約は加えない。これは、4.2.3 節の表 4.7, 4.8 で見たように、文種類によっては複数の文対応を持つ場合が一般的なためである*7。

以下、まず最初に、文対応の推定に先立って文種類 $\hat{\mathbf{t}}^{\text{org}}, \hat{\mathbf{t}}^{\text{rep}}$ を推定し (4.4.2 節)、推定した文種類を素性に投入することで、文種類を考慮した文対応推定を行う方法を示す (4.4.3, 4.4.4 節)。この場合、文種類の推定は文 \mathbf{x}^{org} (又は \mathbf{x}^{rep}) を観測変数、文種類 \mathbf{t}^{org} (又は \mathbf{t}^{rep}) を隠れ変数とした推定問題と考えることができる。続いて、二文書 $\mathbf{x}^{\text{org}}, \mathbf{x}^{\text{rep}}$ から文種類 $\hat{\mathbf{t}}^{\text{org}}, \hat{\mathbf{t}}^{\text{rep}}$ と文対応 \mathbf{y} を同時に推定する方法を示す (4.4.5 節)。

4.4.2 Linear-chain CRF による文種類推定

各文の種類を推定する問題は、系列ラベリング問題とみなすことができる。系列ラベリング問題とは、ある観測された系列 $\mathbf{x} = \{x_1, \dots, x_n\}$ に対し、隠れ変数列 $\mathbf{y} = \{y_1, \dots, y_n\}$ を推定する問題である (浅原, 2007)。文種類推定問題の場合、文列 \mathbf{x} が観測変数、文種類列 \mathbf{t} が隠れ変数に相当する。

自然言語処理分野の系列ラベリング問題に対しては、Linear-chain CRF が広く適用されている。具体的な適用例としては、品詞タグ付け (Lafferty et al., 2001)、テキストチャンキング (Sha & Pereira, 2003)、形態素解析 (Kudo et al., 2004; 工藤他, 2004)、固有表現抽出 (McCallum & Li, 2003; 齋藤他, 2007; 南他, 2011) などがある。例えば、品詞タグ付け問題に Linear-chain CRF を適用すると、各品詞の推定において品詞間の遷移が考慮されるため、学習データから得られた品詞遷移の傾向に基づいて品詞列を付与で

*7 実際に、文対応を一対一に制限した場合の性能を見積もるため、次のような予備実験を行った。まず、多対多関係を許して文対応確率を推定する。次に、文対応確率 $P(y_{i,j} = 1 | \mathbf{x})$ の高い順に、一対多 (多対一) とならない対応のみを確定させる操作を繰り返し行った。結果、多対多関係を許す場合は適合率 46.60%、再現率 60.99% であったのに対し (表 4.12 の combine)、一対一関係に制限した場合は適合率 61.49%、再現率 21.79% と、再現率が大きく低下した。

きる。品詞間の遷移には、例えば助詞が連続することは稀であるなどの特徴が存在するため、これらの特徴を踏まえた品詞付与が実現する。

本研究が対象とする文種類についても、4.2.3 節の表 4.5, 4.6 で見たように遷移傾向が存在する。そこで、本研究ではレビュー文種類・応答文書種類推定のそれぞれについて、Linear-chain CRF を適用することで文種類推定を実現する。Linear-chain CRF による文種類推定は、式 (4.20) の $q_i(\mathbf{x}, y_j, j_{j+1})$ 及び図 4.5 の各 q_i が示すように、隣接するレビュー文種類 t^o_i, t^o_{i+1} (応答文種類 t^r_i, t^r_{i+1}) を依存関係におくため、先に見たような遷移傾向の考慮が実現できる。

4.4.3 Linear-chain CRF による文対応推定

各文間の文対応を推定する問題も、文 \mathbf{x} を観測変数、文対応関係 \mathbf{y} を隠れ変数としたラベリング問題と考えることができる。文対応推定では、予め文種類推定器により推定したレビュー文種類 \hat{t}^{org} 、応答文種類 \hat{t}^{rep} を素性として投入することで、文種類を考慮した推定を実現する。

文対応についても、文種類と同様 4.2.3 節の表 4.10, 4.11 で見たように遷移傾向が存在する。そこで、往信文 x^o_i 又は返信文 x^r_j のいずれかを固定し、Linear-chain CRF を適用することで文対応推定を行う。以下、まず返信文を x^r_k に固定し、各応信文 $x^o_1, x^o_2, \dots, x^o_N$ との文対応 $y_{\bullet, k}$ ($1 \leq \bullet \leq N$) の推定を Linear-chain CRF により行う方法を示す。次に、応信文を x^o_k に固定し、各返信文 $x^r_1, x^r_2, \dots, x^r_N$ との文対応 $y_{k, \bullet}$ ($1 \leq \bullet \leq M$) の推定を Linear-chain CRF により行う方法を示す。

まず、返信文を固定し、各応信文との対応を推定する手順を図 4.7a に示す。この手法では、最初に返信文を x^r_1 に固定し、 $y_{\bullet, 1}$ ($1 \leq \bullet \leq N$) の推定に Linear-chain CRF を適用する。この Linear-chain CRF の場合、 $y_{i, 1}, y_{i-1, 1}$ を依存関係におくため、表 4.10 に示すレビュー文基準での文対応遷移傾向を考慮できる。この推定を、固定する返信文を x^r_1, \dots, x^r_N へと順次移すことで、全体の文対応推定が完了する。

次に、応信文を固定し、各返信文との対応を推定する手順を図 4.7b に示す。この場合は、固定する応信文を x^o_k ($1 \leq k \leq N$) へと順次移し、 $y_{k, \bullet}$ ($1 \leq \bullet \leq M$) の推定に Linear-chain CRF を適用することで、全体の文対応推定を行う。この Linear-chain CRF では、 $y_{k, j}, y_{k, j-1}$ を依存関係におくため、表 4.11 に示す応答文基準での文対応遷移傾向を考慮できる。

以下、図 4.7a に示す往信文間における対応の遷移性を考慮する (返信文を固定し、 $y_{\bullet, j}$ を順次推定する) 手法を L-CRF_{org}、図 4.7b に示す返信文間における対応の遷移性を考慮する (往信文を固定し、 $y_{i, \bullet}$ を順次推定する) 手法を L-CRF_{rep} とする。

4.4.4 2D CRF による文対応推定

次に、2D CRF を適用して文対応を推定する手法について説明する。Linear-chain CRF により文対応を推定する L-CRF_{org} 及び L-CRF_{rep} は、それぞれ往信文間・返信文間の片方のみの遷移性を考慮する手法である。これに対し、2D CRF による文対応推定は、往信文間・返信文間の双方の遷移性を考慮することができる。

2D CRF による推定は、文対応 \mathbf{y} は 2 次元構造を持つため、文 \mathbf{x} を観測変数、文対応 \mathbf{y} を隠れ変数として 2D CRF を適用することで実現できる。この場合の接続性の考慮について、2D CRF の因子グラフ (図 4.6) を用いて説明する。Linear-chain CRF を繰り返し適用する手法 L-CRF_{org} 及び L-CRF_{rep} は、図 4.6 に示す因子のうち各 y を縦方向 $y_{i-1, j}, y_{i, j}$ 又は横方向 $y_{i, j}, y_{i, j-1}$ を結ぶ因子のみを残した場合に相当し、各列 $y_{\bullet, i}$ または各行 $y_{i, \bullet}$ ごとに推定を行う。これに対し、2D CRF では各 y が縦・横の両方向

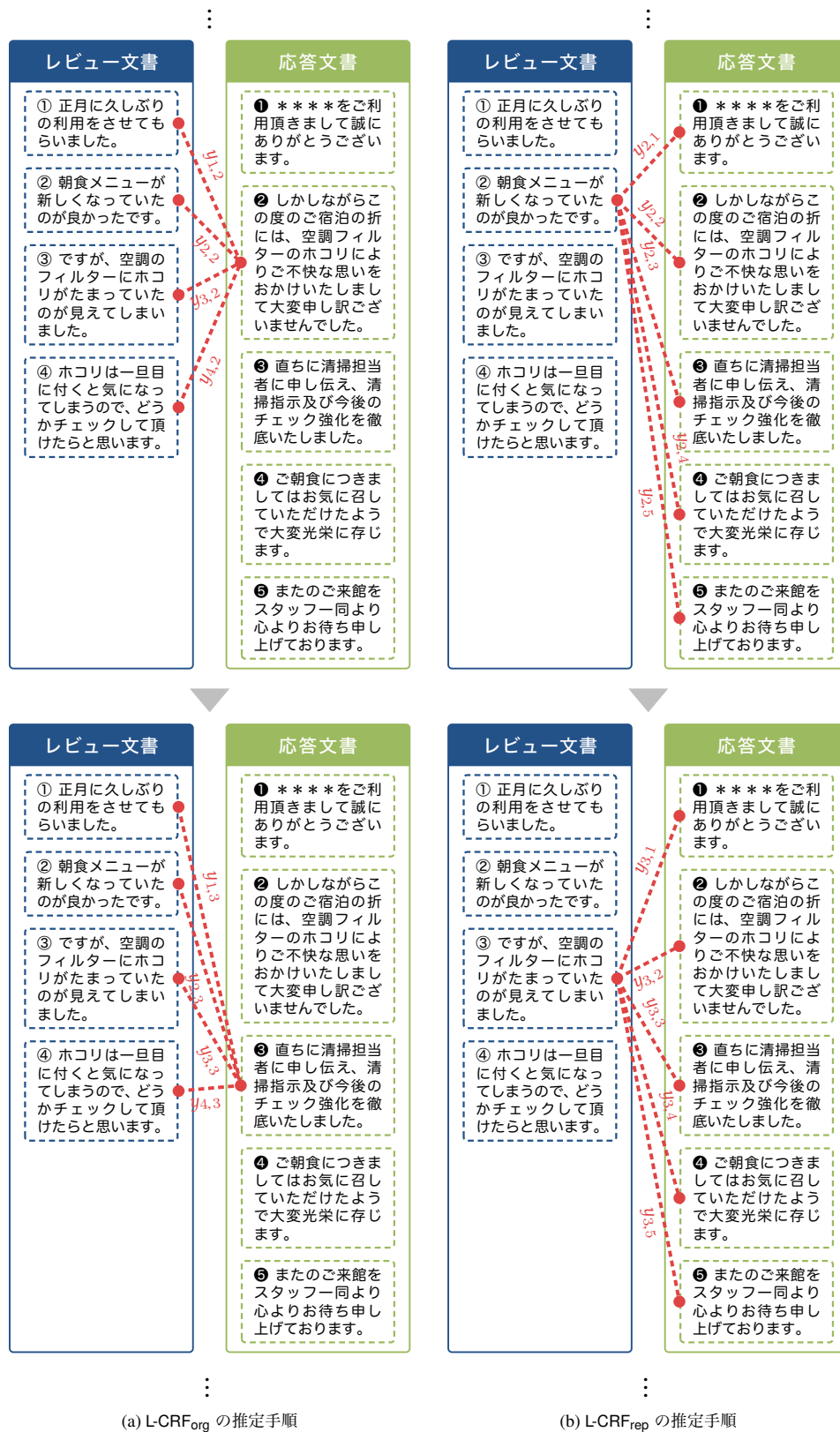


図 4.7: Linear-chain CRF を用いた文対応の推定手順

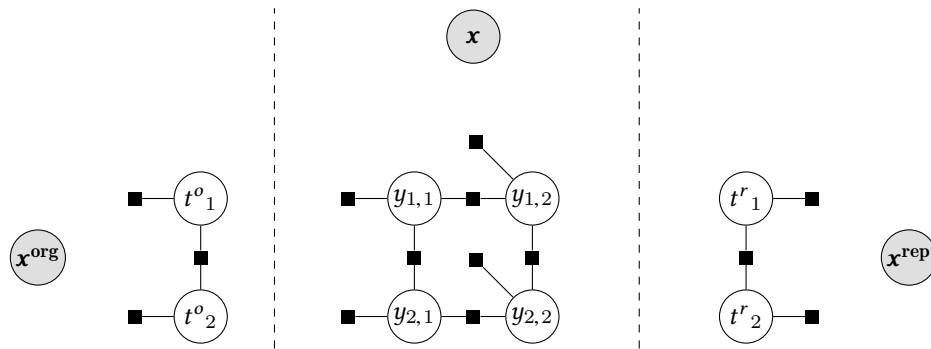


図 4.8: 統合前のモデル (往信文・返信文が共に 2 文の場合。x から各因子への接続は省略している)。左から順に、往信文種類モデル、文対応モデル (2D CRF)、返信文種類モデル

で結ばれているため、往信文間及び返信文間における双方の対応の接続性 (表 4.10, 4.11 で示した傾向) を考慮するモデルとなる。すなわち、 $L\text{-CRF}_{\text{org}}$ 及び $L\text{-CRF}_{\text{rep}}$ の双方の性質を併せ持つモデルと言える。以下、2D CRF による文対応推定手法を 2D CRF と呼ぶ。

4.4.5 文種類・文対応推定モデルの統合

ここまでに説明した文対応推定 $L\text{-CRF}_{\text{org}}$, $L\text{-CRF}_{\text{rep}}$, 2D CRF は、いずれも予め推定した文種類情報を文対応推定の素性として用いることにより文種類を考慮した文対応の推定を行う。これらの推定手法は、文種類と文対応の推定を別のステップで行うため、文種類と文対応の間に独立性を仮定していることになる。しかし、4.2.3 節で見たように、文種類と文対応の間にはある程度のある関係があることが分かっている。また、文対応推定時に投入する文種類はいくらかの誤りが含まれる可能性が高いため、文種類推定時の誤りがそのまま文対応推定に影響を与える問題もある。そこで、本研究では文種類と文対応を推定するモデルを統合し、両者を同時に推定するモデル *combine* を提案する。

本研究では統合の基本となるモデルとして、文種類推定に Linear-chain CRF を用いた往信文種類モデル・返信文種類モデル (4.4.2 節)、文対応推定に 2D CRF を用いた文対応モデル (4.4.4 節) を考える。ここで文種類推定に Linear-chain CRF による推定モデルに用いる理由は、4.4.2 節で説明したように Linear-chain CRF は隠れ変数間に遷移傾向がある場合の推定に適している推定モデルなためである。また、文対応推定に 2D CRF による推定モデルを用いる理由は、4.4.4 節で説明したように往信文間及び返信文間における双方の遷移傾向を考慮できる推定モデルなためである。

これらのモデルに対し、文種類変数と文対応変数に依存する因子関数を新たに加えることで、統合モデルを実現する。依存関係を全て考慮するモデルは、観測変数 \mathbf{x}^{org} , \mathbf{x}^{rep} 、文種類 t^{org} , t^{rep} 及び文対応 \mathbf{y} を全て依存関係におくモデルであるが、このようなモデルはスパースネス問題の影響を強く受けしてしまう。基本となる Linear-chain CRF 及び 2D CRF は、スパースネス問題を回避するために少ないながらも比較的效果が高い依存関係を仮定していることから、統合モデルに加える因子関数も、関係する文種類・文対応を依存関係におくことができる最小限の関数とする。以下、具体的な統合方法について因子グラフを用いながら説明する。

まず最初に、統合の元となる各モデルの因子グラフを図 4.8 に示す。ここで、各モデルの因子グラフ構造を 4.3.2 節の記法を用いて記述すると、以下の通りである。

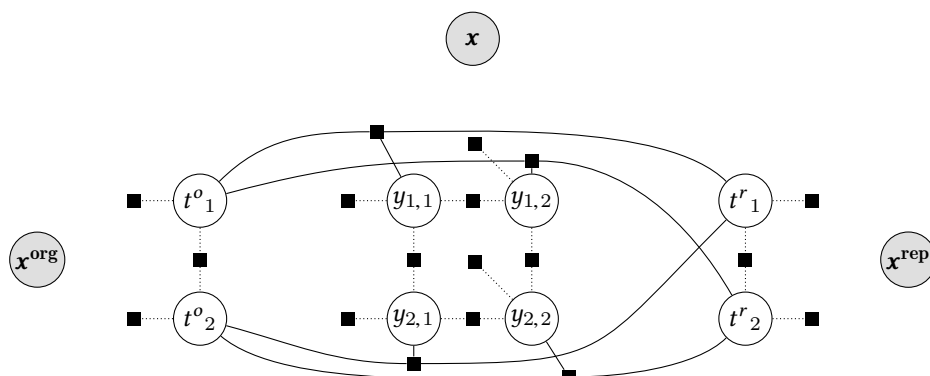


図 4.9: 統合後のモデル (combine; 往信文・返信文が共に 2 文の場合。x から各因子への接続は省略している)。点線は統合前から存在した因子を結ぶリンク、実線は新たに加えた因子を結ぶリンク

往信文種類モデル

Linear-chain CRF により往信文種類を推定する。観測変数は往信文 \mathbf{x}^{org} 、隠れ変数は往信文種類 t^{org} である。観測変数と隠れ変数を結ぶ因子を f^o_i 、隠れ変数同士を結ぶ因子を g^o_j とする ($1 \leq i \leq N, 1 \leq j \leq N-1$)。モデルの因子グラフを $\mathcal{G}_{\text{otype}} = \{\mathcal{V}_{\text{otype}}, \mathcal{F}_{\text{otype}}, \mathcal{E}_{\text{otype}}\}$ とすると、 $\mathcal{V}_{\text{otype}}, \mathcal{F}_{\text{otype}}, \mathcal{E}_{\text{otype}}$ はそれぞれ以下の通りである。因子グラフを図 4.8 (左) に示す。

$$\mathcal{V}_{\text{otype}} = \{\mathbf{x}^{\text{org}}\} \cup \{t^o_i\}_{i=1}^N \quad (4.24)$$

$$\mathcal{F}_{\text{otype}} = \{f^o_i\}_{i=1}^N \cup \{g^o_j\}_{j=1}^{N-1} \quad (4.25)$$

$$\mathcal{E}_{\text{otype}} = \{(f^o_i, \mathbf{x}^{\text{org}})\}_{i=1}^N \cup \{(f^o_i, t^o_i)\}_{i=1}^N \\ \cup \{(g^o_j, \mathbf{x}^{\text{org}})\}_{j=1}^{N-1} \cup \{(g^o_j, t^o_j)\}_{j=1}^{N-1} \cup \{(g^o_j, t^o_{j+1})\}_{j=1}^{N-1} \quad (4.26)$$

返信文種類モデル

Linear-chain CRF により返信文種類を推定する。観測変数は返信文 \mathbf{x}^{rep} 、隠れ変数は返信文種類 t^{rep} である。観測変数と隠れ変数を結ぶ因子を f^r_i 、隠れ変数同士を結ぶ因子を g^r_j とする ($1 \leq i \leq M, 1 \leq j \leq M-1$)。モデルの因子グラフを $\mathcal{G}_{\text{rtype}} = \{\mathcal{V}_{\text{rtype}}, \mathcal{F}_{\text{rtype}}, \mathcal{E}_{\text{rtype}}\}$ とすると、 $\mathcal{V}_{\text{rtype}}, \mathcal{F}_{\text{rtype}}, \mathcal{E}_{\text{rtype}}$ はそれぞれ以下の通りである。因子グラフを図 4.8 (右) に示す。

$$\mathcal{V}_{\text{rtype}} = \{\mathbf{x}^{\text{rep}}\} \cup \{t^r_i\}_{i=1}^M \quad (4.27)$$

$$\mathcal{F}_{\text{rtype}} = \{f^r_i\}_{i=1}^M \cup \{g^r_j\}_{j=1}^{M-1} \quad (4.28)$$

$$\mathcal{E}_{\text{rtype}} = \{(f^r_i, \mathbf{x}^{\text{rep}})\}_{i=1}^M \cup \{(f^r_i, t^r_i)\}_{i=1}^M \\ \cup \{(g^r_j, \mathbf{x}^{\text{rep}})\}_{j=1}^{M-1} \cup \{(g^r_j, t^r_j)\}_{j=1}^{M-1} \cup \{(g^r_j, t^r_{j+1})\}_{j=1}^{M-1} \quad (4.29)$$

文対応モデル (2D CRF)

2D CRF により文対応を推定する。観測変数は往信文 \mathbf{x}^{org} 及び返信文 \mathbf{x}^{rep} (これらをまとめて \mathbf{x} とする)、隠れ変数は文対応 \mathbf{y} である。観測変数と隠れ変数を結ぶ因子を $f_{i,j}$ 、隠れ変数同士を結ぶ因子のうち、往信文を固定して隣接する返信文への対応間を考慮する ($y_{i,j}$ を横に結ぶ) 因子を $g_{k,l}$ 、返信文を固定して隣接する往信文からの対応間を考慮する ($y_{i,j}$ を縦に結ぶ) 因子を $h_{k,l}$ とする ($1 \leq i \leq N, 1 \leq j \leq M, 1 \leq k \leq N-1, 1 \leq l \leq M-1$)。因子グラフを図 4.8 (中央) に示す。モデルの因子グラフを $\mathcal{G}_{\text{relation}} = \{\mathcal{V}_{\text{relation}}, \mathcal{F}_{\text{relation}}, \mathcal{E}_{\text{relation}}\}$ とすると、

$\mathcal{V}_{\text{relation}}, \mathcal{F}_{\text{relation}}, \mathcal{E}_{\text{relation}}$ はそれぞれ以下の通りである。

$$\mathcal{V}_{\text{relation}} = \{\mathbf{x}\} \cup \{y_{i,j}\}_{i=1,j=1}^{i=N,j=M} \quad (4.30)$$

$$\mathcal{F}_{\text{relation}} = \{f_{i,j}\}_{i=1,j=1}^{i=N,j=M} \cup \{g_{k,l}\}_{k=1,l=1}^{k=N-1,l=M-1} \cup \{h_{k,l}\}_{k=1,l=1}^{k=N-1,l=M-1} \quad (4.31)$$

$$\begin{aligned} \mathcal{E}_{\text{relation}} = & \{(f_{i,j}, \mathbf{x})\}_{i=1,j=1}^{i=N,j=M} \cup \{(f_{i,j}, y_{i,j})\}_{i=1,j=1}^{i=N,j=M} \\ & \cup \{(g_{k,l}, \mathbf{x})\}_{k=1,l=1}^{k=N-1,l=M-1} \cup \{(h_{k,l}, \mathbf{x})\}_{k=1,l=1}^{k=N-1,l=M-1} \\ & \cup \{(g_{k,l}, y_{k,l})\}_{k=1,l=1}^{k=N-1,l=M-1} \cup \{(g_{k,l}, y_{k,l+1})\}_{k=1,l=1}^{k=N-1,l=M-1} \\ & \cup \{(h_{k,l}, y_{k,l})\}_{k=1,l=1}^{k=N-1,l=M-1} \cup \{(h_{k,l}, y_{k+1,l})\}_{k=1,l=1}^{k=N-1,l=M-1} \end{aligned} \quad (4.32)$$

次に、以上の3モデルを統合した提案モデル **combine** について説明する。**combine** モデルでは、関係する文種類・文対応を依存関係におく因子関数として $F_{i,j}$ ($1 \leq i \leq N, 1 \leq j \leq M$) を新たに加える。因子関数 $F_{i,j}$ は、 i 番目の往信文 x^o_i と j 番目の返信文 x^r_j と関係する、往信文種類 t^o_i 、返信文種類 t^r_j 、文対応 $y_{i,j}$ を依存関係におく因子関数である。この因子関数は、 i 番目の往信文と j 番目の返信文に注目した際に直接関わる変数は $t^o_i, t^r_j, y_{i,j}$ であることから、関係する文種類・文対応変数を依存関係におくことができる最小限の因子関数である。これにより、文種類・文対応変数を非独立的に扱った文種類・文対応の同時推定が可能になる。提案手法全体のグラフ構造を図 4.9 に示す。因子グラフ構造を具体的な式で記述すると、以下の通りである。以下、必要に応じて統合したモデルを **combine** と呼ぶ。

統合モデル (combine)

観測変数は往信文 \mathbf{x}^{org} 及び返信文 \mathbf{x}^{rep} (これらをまとめて \mathbf{x} とする)、隠れ変数は往信文種類 t^{org} ・返信文種類 t^{rep} ・文対応 \mathbf{y} である。新たに導入する因子を $F_{i,j}$ とする ($1 \leq i \leq N, 1 \leq j \leq M$)。その他の因子は統合前の各モデルと同様である。

モデルの因子グラフを $\mathcal{G}_{\text{combine}} = \{\mathcal{V}_{\text{combine}}, \mathcal{F}_{\text{combine}}, \mathcal{E}_{\text{combine}}\}$ とすると、

$\mathcal{V}_{\text{combine}}, \mathcal{F}_{\text{combine}}, \mathcal{E}_{\text{combine}}$ はそれぞれ以下の通りである (新たに加わった因子に関わる部分を下線によって強調している)。因子グラフを図 4.9 に示す。

$$\mathcal{V}_{\text{combine}} = \{\mathbf{x}\} \cup \{t^o_i\}_{i=1}^N \cup \{t^r_j\}_{j=1}^M \cup \{y_{i,j}\}_{i=1,j=1}^{i=N,j=M} \quad (4.33)$$

$$\mathcal{F}_{\text{combine}} = \mathcal{F}_{\text{otype}} \cup \mathcal{F}_{\text{rtype}} \cup \mathcal{F}_{\text{relation}} \cup \underline{\{F_{i,j}\}_{i=1,j=1}^{i=N,j=M}} \quad (4.34)$$

$$\begin{aligned} \mathcal{E}_{\text{combine}} = & \mathcal{E}_{\text{otype}} \cup \mathcal{E}_{\text{rtype}} \cup \mathcal{E}_{\text{relation}} \\ & \cup \underline{\{(F_{i,j}, \mathbf{x})\}_{i=1,j=1}^{i=N,j=M}} \\ & \cup \underline{\{(F_{i,j}, t^o_i)\}_{i=1,j=1}^{i=N,j=M} \cup \{(F_{i,j}, t^r_j)\}_{i=1,j=1}^{i=N,j=M} \cup \{(F_{i,j}, y_{i,j})\}_{i=1,j=1}^{i=N,j=M}} \end{aligned} \quad (4.35)$$

なお、Linear-chain CRF などでは各隠れ変数の周辺確率 $P(y|\mathbf{x})$ を Forward-Backward アルゴリズムにより効率的に求めることができるが、統合モデル **combine** のように閉路を含む因子グラフ上の CRF に適用することができない。そのため、Tree Based reParameterization (TRP) (Wainwright et al., 2001) などにより近似的に求める必要がある。

4.5 評価実験

4.5.1 実験条件

実験に用いるデータは、4.2.3 節で用いた楽天トラベルのレビュー文書・応答文書と同じデータを用いる。4.2.3 節で得られた 1,000 文書対に対し、5 分割交差検定を適用して評価を行う。

実験で比較する手法は次の5つである。まず、4.4.3, 4.4.4節で説明した L-CRF_{org}, L-CRF_{rep}, 2D CRF の3種類を用いる。また、4.4.5節で説明した統合モデル combine を用いる。加えて、系列ラベリング問題ではなく二値分類問題と考えるモデルとしてロジスティック回帰 (logistic) でも性能を調査する。

CRFの各モデルのパラメータ学習・利用には MALLET 2.0.7 (McCallum, 2002) 中の GRMM (Sutton, 2006) を用いた。GRMM に用いたパラメータはデフォルト (TRP の最大 iteration 回数 1,000 回、TRP の収束判定用の値 0.01) とし、周辺確率の計算には TRP (Wainwright et al., 2001) を利用した。なお、GRMM は CRF 学習パラメータの正則化に L2 正則化 (Chen & Rosenfeld, 1999) を利用している。また、ロジスティック回帰のパラメータ学習・利用には scikit-learn 0.15.1 (Pedregosa et al., 2011) を用い、正則化には L2 正則化を利用した。

文種類の推定には、文を構成する単語 (表層形) を素性として用いる。また、文対応の推定、及び combine モデルにおいて新たに追加した因子には以下の素性を用いる。なお、レビュー文・応答文の単語は、ベクトル空間モデルによりベクトル化して投入する。

- レビュー文を構成する単語
- 応答文を構成する単語
- レビュー文・応答文をベクトル空間モデルによりベクトル化した際のコサイン類似度 (0 ~ 1 の値)
- 予め文種類モデルで推定したレビュー文・応答文の種類 (combine モデル以外^{*8})

文種類・文対応の推定に単語を素性として用いる理由は、文種類・文対応いずれの推定においても文の表層は最も重要なためである。ただし、文をそのままの形で扱くと、自然言語文の大半はユニークであることから正しく推定が行えないと言う問題がある。そこで、文を単語のレベルに分解することで、より汎用的な学習を行う。また、文対応推定にコサイン類似度を用いた理由は、4.2.3節の図 4.2, 4.3 で見たように、文対応が存在する文対の方が存在しない文対よりもコサイン類似度が高いと言う傾向があり、この傾向を推定に利用するためである。なお、単語分割には MeCab 0.994 (Kudo et al., 2004) を利用した。また、素性として利用する単語からは予め stop-word を除去しており、1,000 文書対全体では 9,300 種類の単語が存在した。更に、文種類を素性として用いる理由は、4.2.3節で見たような、文種類と文対応の間の関係を推定に利用するためである^{*9}

文対応推定性能の評価は、適合率 (Precision) ・再現率 (Recall)、及びそれらの調和平均である F 値から行う。すなわち、考えうる全ての文対応の可能性から正しい文対応を探す課題とみなし、次の式で計算する。

$$\text{Recall} = \frac{\text{正しく推定できた対応数}}{\text{評価データ中に存在する文対応数}}$$

$$\text{Precision} = \frac{\text{正しく推定できた文対応数}}{\text{システムが文対応有りとして出力した文対応数}}$$

本実験では、以下に説明する手法により適合率・再現率を調整し、これらの性能がどのように変化するかを調査する。具体的には、文同士が対応する確率 $P(y_{i,j} = 1)$ と対応しない確率 $P(y_{i,j} = 0)$ の比率を

^{*8} combine モデルの場合は、レビュー文種類・応答文種類は文対応と同時に推定されるため、これらを陽に素性として追加する必要はない。combine モデルにおいては文種類と文対応を同時に考慮する因子が存在するため、文種類を考慮した文対応推定が実現できる。

^{*9} 文種類情報の有用性を確認するため、文種類を素性として利用しない場合の性能を調査する予備実験を行った。結果、文種類を利用しない場合 L-CRF_{org}, L-CRF_{rep}, 2D CRF いずれにおいても適合率が3ポイント程度、再現率が1ポイント程度低下した。このため、文対応推定において文種類情報は有用であると言える。

表 4.12: 文対応推定性能 (分割交差検定での F 値最大点におけるマイクロ平均値)

手法	適合率	再現率	F 値	閾値 α
logistic	41.99%	64.00%	50.70%	-0.57
L-CRF _{org}	42.65%	60.65%	50.08%	-1.37
L-CRF _{rep}	42.94%	57.58%	49.20%	-1.10
2D CRF	44.11%	60.45%	51.00%	-0.87
combine	46.60%	60.99%	52.83%	-0.97

取り、閾値を与えて閾値以上か否かで文対応の有無の出力を変更する。すなわち、文対応 $y_{i,j}$ の最終的な出力 $\hat{y}_{i,j}$ は閾値 α を用いて次の式 (4.36) のようにする。

$$\hat{y}_{ij} = \begin{cases} 1 & \text{if } \log \frac{P(y_{i,j}=1|\mathbf{x})}{P(y_{i,j}=0|\mathbf{x})} > \alpha \\ 0 & \text{otherwise} \end{cases} \quad (4.36)$$

4.5.2 実験結果と考察

まず、文対応推定の適合率・再現率を表 4.12 に示す。各数値は適合率・再現率を調整した際に学習データにおいて F 値が最大となる点を用い、5 分割交差検定でのマイクロ平均値を計算した数値である。なお、表 4.12 に示す結果の F 値に対してブートストラップ検定で得られた p 値を Holm 法 (Holm, 1979) によって調整した有意水準と比較することで多重比較を行い、統合モデル combine は他手法全てに対して統計的有意差があることを確認している (有意水準 5%)^{*10}。また、combine モデルの F 値最大点における文種類ごとの推定性能を表 4.13, 4.14 に示す^{*11}。

表 4.12 を見ると、統合モデル combine の絶対的な性能は適合率 46.60%、再現率 60.99% であった。この性能を本研究課題の応用と言う観点から見ると、例えば本章冒頭で述べた応用例である「(1) 未対応文の検出による返信文書作成の支援」のためには、検出漏れが無いようにより高い再現率が必要である。特に、表 4.13 に示す〈ネガティブ感想〉や〈要求・要望〉は、第 1 章で述べたようにサービス失敗への顧客対応は重要であるため、特に高い再現率であることが望ましい。また、「(2) 定型的返信文の自動生成」「(3) 非定型的返信文の返答例提示」のためには、自動生成・半自動生成の学習データに投入できる程度の質を確保するためより高い適合率が必要である。相対的な性能では、F 値最大点において統合モデル combine は他手法よりも高い性能であったが、各応用のためにはより高い性能が必要であると考えられる。これに対しては、本研究の問題設定はいずれの応用も視野に入れた設定であるため、目的とする応用に合わせた改変を加えることである程度の対処が可能と考えている。例えば、「(1) 未対応文の検出による返信文書作成の支援」の場合は、特に重要な〈ネガティブ感想〉や〈要求・要望〉などの特定文種類に特化し、再現率を重視するなどの処理を加えるなどの方法がある。

次に、統合モデル combine の文種類ごとの推定性能に注目する (表 4.13, 4.14)。いずれの文書を基準にしても文種類によって性能に大きな差があるが、この主な理由は学習データ量の差によるものと考えている。すなわち、一部の文種類はほとんど文対応を持たないことや、そもそも当該文種類を持つ文の

^{*10} ブートストラップ検定におけるブートストラップ回数は 1,000 回とした。

^{*11} 表 4.14 で全ての項目が「-」となっている文種類 (〈結びでの感謝〉〈署名・フッター〉) は今回のデータには該当する文種類に文対応がなかったもの、F 値が「-」となっている文種類 (〈定型的挨拶〉) は推定結果が全て false-positive であったものである。

表 4.13: 提案モデル combine におけるレビュー文種類ごとの文対応推定性能 (F 値最大点)

レビュー文種類	適合率	再現率	F 値
宿泊説明	26.02%	62.50%	36.74%
感謝・応援	5.95%	36.36%	10.22%
プラン名	6.45%	100.00%	12.12%
再泊・推薦意向	14.52%	36.00%	20.69%
不再泊・不薦意向	6.20%	100.00%	11.68%
ポジティブ感想	48.01%	68.69%	56.52%
ポジ/ネガ感想	59.93%	60.13%	60.03%
ネガティブ感想	58.35%	56.13%	57.22%
ニュートラル感想	12.28%	46.05%	19.39%
要求・要望	50.68%	52.71%	51.68%
情報追加	12.50%	42.71%	19.34%
個別事情の説明	22.28%	40.88%	28.84%
レビュー文全体	46.60%	60.99%	52.83%

表 4.14: 提案モデル combine における応答文種類ごとの文対応推定性能 (F 値最大点)

応答文種類	適合率	再現率	F 値
宿泊御礼	30.12%	64.66%	41.10%
投稿御礼	1.69%	13.46%	3.00%
結びでの感謝	-	-	-
心掛け・決意	14.61%	30.53%	19.76%
再泊願い	5.60%	19.29%	8.68%
署名・フッター	-	-	-
定型的挨拶	0.00%	0.00%	-
ほめへの感謝	59.93%	77.19%	67.47%
お詫び	50.08%	60.37%	54.74%
恥じ入り	42.68%	43.75%	43.21%
提案・指摘感謝	46.73%	35.48%	40.33%
事実述べ	52.37%	63.68%	57.47%
具体的対応明示	47.38%	63.96%	54.43%
抽象的対応明示	52.42%	46.14%	49.08%
具体的検討明示	60.36%	59.30%	59.82%
抽象的検討明示	53.80%	53.40%	53.60%
対処明示	45.97%	51.33%	48.50%
了承願い	61.43%	69.35%	65.15%
対話	26.14%	33.87%	29.51%
情報追加	38.32%	52.02%	44.13%
応答文全体	46.60%	60.99%	52.83%

表 4.15: 2D CRF と combine 間における出力の変
化 (正解: 文対応あり)

		2D CRF	
		TP	FN
combine	TP	2,227	513
	FN	488	1,264

表 4.16: 2D CRF と combine 間における出力の変
化 (正解: 文対応なし)

		2D CRF	
		TN	FP
combine	TN	21,644	1,414
	FP	1,112	2,027

出現回数が少ないことに起因して学習が難しくなっている。特に前者については、例えばレビュー文種類<感謝・応援><プラン名>や応答文種類<投稿御礼>について表 4.7, 4.8 を見ると、登場する回数が多いものの文対応を持つものは極めて少ないことが分かる。これらの文種類については再現率が高いものの適合率が極めて低いことから、過剰に対応有りと推定してしまっていることが分かる。このように対応する可能性が低い文種類については、推定後に予め人手で作成したルールによりフィルタリングするなどにより解決できると考えている。

更に、統合前モデル 2D CRF と統合後モデル combine の間における出力の変化に注目する。表 4.15, 4.16 に、同じ文対応について 2D CRF と combine による推定結果が一致した・しなかった数を示す。表

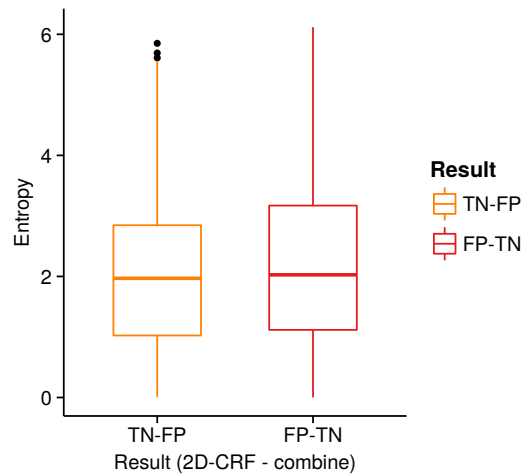


図 4.10: 2D CRF と combine 間において出力が TN → FP 又は FP → TN と変化した文対応に対応する文種類のエントロピー

4.15 は正解データが文対応ありのもの、表 4.16 は文対応なしのものである。表 4.15 の TP, FP はそれぞれ True-Positive, False-Negative を、表 4.16 の TN, FP は True-Negative, False-Positive を表す。いずれの表とも、対角の値は 2D CRF と combine 間で推定結果が一致した数を表し、非対角成分は一致しなかった数を表す。2D CRF と combine の性能差は推定結果が一致していないものから生じるため、非対角成分（下線数値）に注目すると、表 4.15 では大きな違いがなく、表 4.16 では 2D CRF が False-Positive であったものを combine で True-Negative と判断した数が、combine が True-Negative であったものを 2D CRF で False-Positive と判断した数が上回っていることが分かる。すなわち、2D CRF で過剰に対応ありとしていたものを、combine では対応なしとすることにより、同程度の再現率を保ちながら適合率を高めている。これらの文対応の特徴を調べるため、文対応と対応するレビュー文・応答文について、Linear-chain CRF による文種類確率の推定結果のエントロピーを調査した。レビュー文・応答文種類確率のエントロピーの和について、出力の変化から見た分布をボックスプロットに取った図を図 4.10 に示す。図 4.10 は、2D CRF で True-Negative であり combine で False-Positive であった二文 (TN-FP) より、2D CRF で False-Positive であり combine で True-Negative であった二文 (FP-TN) の文種類確率分布の方が高いエントロピーを持つ傾向を表している。これは、2D CRF ではエントロピーが高い、すなわち文種類の推定が難しいものでも一意に定めた文種類を用いたために対応ありと出力して誤っている (FP) ものを、combine では対応なしと出力する (TN) ことにより抑制できていることを表す。これは、combine が文種類を一意に定めず、同時に推定することによる効果であると考えている。

続いて、式 (4.36) の閾値 α を変化させて適合率・再現率を調整した際の Precision-Recall 曲線を図 4.11 に示す。図 4.11 より中程度の再現率 (Recall: 0.25 ~ 0.75) でも combine は概ね他手法よりも高い適合率であり、多くの場合において高い性能であったと言える。一方で、低再現率 (Recall: 0.0 ~ 0.25) では、中再現率で高い適合率であった combine よりもロジスティック回帰 logistic や L-CRF_{rep} の方が高い性能であった。この原因を調べるため、低再現率と中再現率においてどのような文対応が推定できているかをコサイン類似度の観点から調査した。ここで、各 Recall 値の性能時において推定できた全ての文対応に対しコサイン類似度を計算し、平均を取った値の変化をグラフにした図を図 4.12 に示す。図 4.12 より、いずれの手法においても低い再現率値においてはコサイン類似度の平均が高く、徐々にコサ

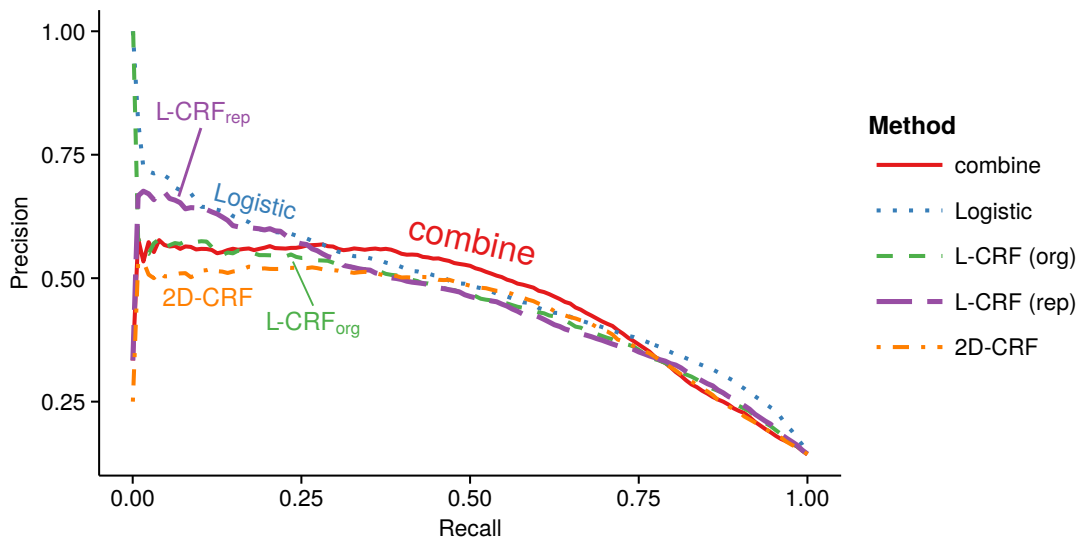


図 4.11: 実験結果の Precision-Recall 曲線

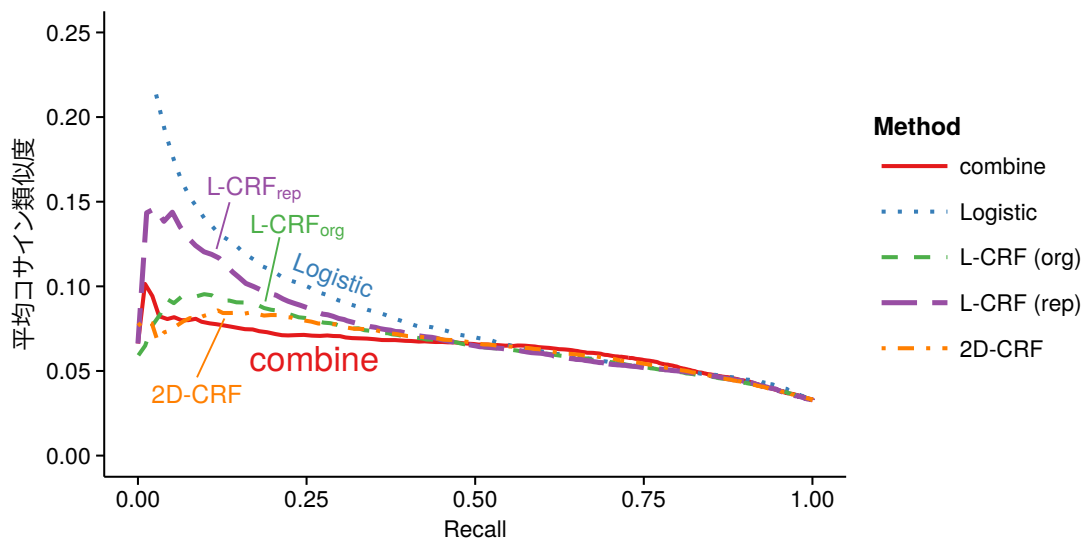


図 4.12: 平均コサイン類似度値-Recall 曲線。各 Recall 値において推定された全ての文対応に対しコサイン類似度を計算し、平均を取った値の変化を示す

イン類似度の平均は下がって行くことが分かる。中でも **combine** は低再現率においてコサイン類似度の平均が他手法に比較すると特に低いことから、コサイン類似度の重要性を低く見ているために高類似度の文間に見られる文対応を見落としていると考えている^{*12}。これに対しては、**combine** では他の手法よりも単純な単語マッチなどでは推定が困難な文対応もある程度発見できるという特徴を持つことでもあ

^{*12} なお、例えば再現率 10% において推定された文対応のコサイン類似度について、**logistic** の分散は 0.052、**combine** の分散は 0.035 であった。特に、コサイン類似度が 0 であった文対の割合は、**logistic** の場合は 12.3% であったのに対し、**combine** の場合は 48.8% と大きな差があった。

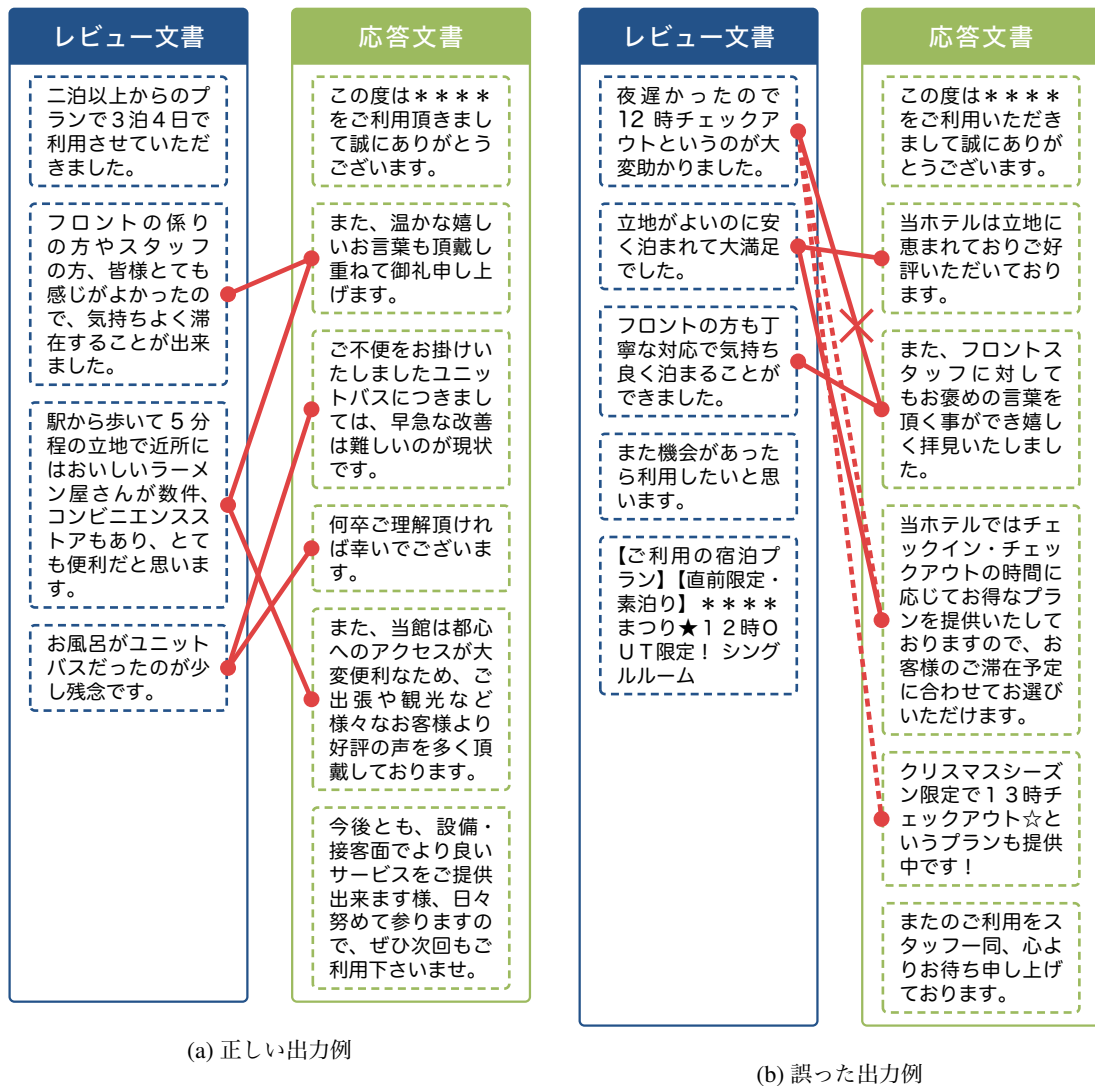


図 4.13: combine モデルによる推定例。実線は対応有りと推定された正しい文対応を示す。実線に × 記号があるものは対応有りと推定されたが実際には対応していないもの、破線は対応無しと推定されたが実際には対応しているものを示す。

るため、コサイン類似度の値によって推定手法を切り替えるなどで様々な文対にも対応可能になると考えている。すなわち、コサイン類似度が高い文対ではロジスティック回帰や $L-CRF_{rep}$ など、低い文対では combine を用いることで、より推定性能が向上すると考えている。

次に、実際の combine モデルの出力例を図 4.13 に示す (表 4.12 に示す F 値最大点における出力結果)。図中の実線が推定によって得られた正しい文対応を示し、実線に × 記号があるものは対応有りと推定されたが実際には対応していないもの、破線は対応無しと推定されたが実際には対応しているものを示す。

図 4.13a は誤りなく文対応を推定できた例である。例えばレビュー文「フロントの係りの方やスタッフの方、皆様とても感じがよかったので、気持ちよく滞在することが出来ました。」「駅から歩いて 5 分程の立地で近所にはおいしいラーメン屋さんが数件、コンビニエンスストアもあり、とても便利だと思います。」お風呂がユニットバスだったのが少し残念です。」はいずれも対応先の応答文「また、温かな嬉しいお言葉も頂戴し重ねて御礼申し上げます。」と共通する内容語が存在しないが、正しく文対応が推定できている。これは、それぞれの文種類を <ポジティブ感想> <ほ

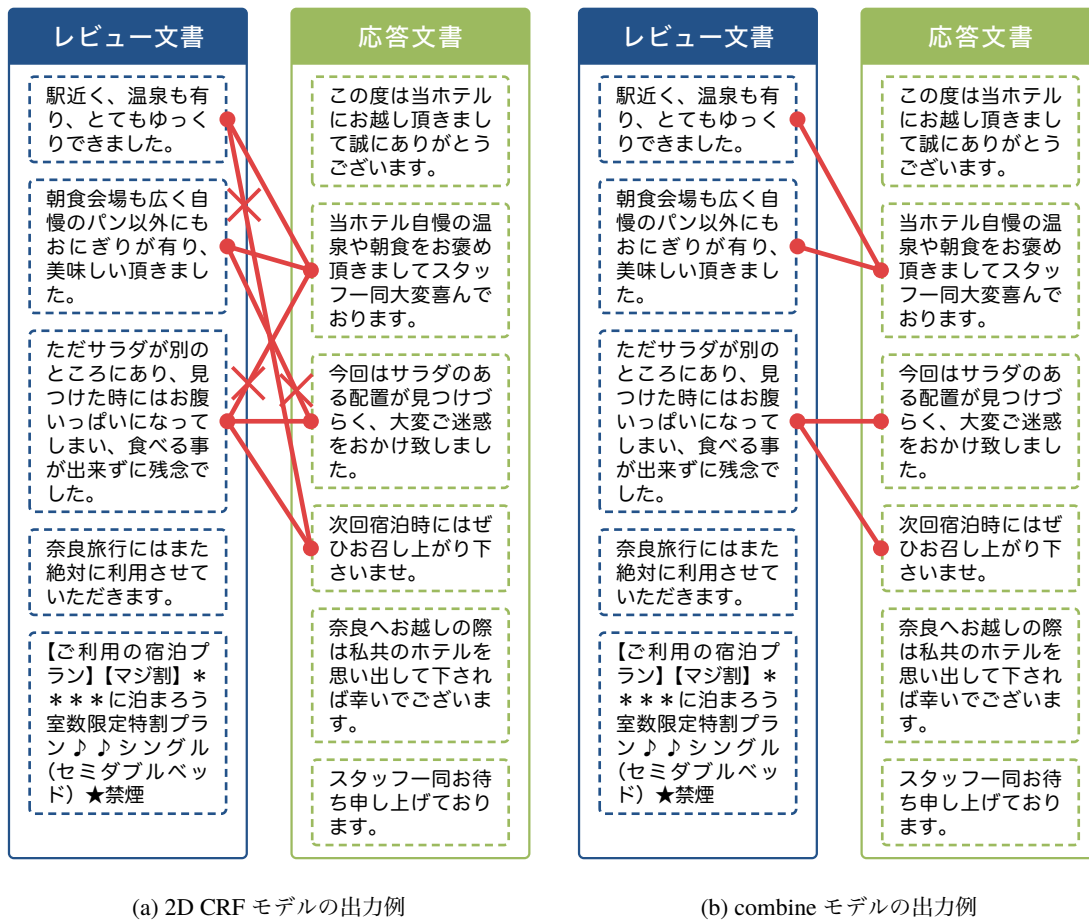


図 4.14: combine モデルと 2D CRF モデルの推定例

めへの感謝> と正しく推定できており、加えて「気持ちよく」「おいしい」「便利」と言った語が現れる文と「御礼」と言った言葉が現れる文の間には対応する可能性が高いと言った傾向をうまく学習できたことによると考えている。

また、図 4.13b は誤った推定が含まれている例である。例えばレビュー文「夜遅かったので…」は応答文「また、フロントスタッフに対しても…」と対応していると推定して誤っている。これは逆に、文種類がそれぞれ<ポジティブ感想><ほめへの感謝>であることや、レビュー文中に「チェックアウト」が、応答文中に「フロント」「スタッフ」などの語が出現すると対応しやすいと言う傾向に影響されているためであると考えている。この場合、それぞれの文で触れられている対象が、チェックアウト時刻そのものなのか、チェックアウト時のスタッフの対応なのかを区別できればより正確な推定が可能になる。同様に、「夜遅かったので…」に対して応答文「当ホテルでは…」 「クリスマスシーズン…」の間の文対応を発見することができなかった問題についても、それぞれの文でチェックアウト時刻に触れられていることを特定できれば対応を発見できる可能性が向上すると考えている。

最後に、combine モデルと 2D CRF モデルによる出力の比較例を図 4.14 に示す (表 4.12 に示す F 値最大点における出力結果)。この例では、2D CRF モデルでは誤って対応ありと出力したペアに対して、combine モデルでは正しく対応がないと出力できている。2D CRF モデルが誤った理由の一つとし

て、文対応推定の前提処理である文種類推定の誤りによる影響があると考えている。この例では、応答文「今回はサラダのある…」及び「次回宿泊時には…」に対して文種類〈情報追加〉が誤って推定されているが（正しくは〈お詫び〉及び〈対話〉）、〈情報追加〉は関連した語が登場するレビュー文と対応を持ちやすいと言う傾向があるため、過剰に対応有りとして出力されていると考えている。これに対し、combine モデルでは文種類と文対応を同時に推定するため、事前の文種類推定における誤りに影響されると言ったことはないため正しく推定できている。これは、表 4.15, 4.16 で見た、「2D CRF で過剰に対応ありとしていたものを、combine では対応なしとする傾向」を示す典型例の一つでもある。

4.6 関連研究

二文書以上の文書間において、文書を跨いだ文同士の関係に踏み込んだ研究は新聞記事を対象としたものが多い (Radev, 2000; 難波他, 2005; 宮部他, 2006, 2005)。Radev (2000) は新聞記事間に観察できる文間関係を「同等 (Equivalence)」「反対 (Contradiction)」などの 24 種類に分類する Cross-Document Structure Theory を提案した。これら文間関係のうち、宮部他 (2006, 2005) は「同等」「推移」関係の特定に、難波他 (2005) は「推移」「更新」関係の特定に特化した自動推定手法を提案している。これらの各研究では「同等」「推移」「更新」関係を特定するために文同士が類似しているなどの特徴を利用して、これらの研究と本研究を比較すると、まず、これらの研究が扱う新聞記事間における文対応と本研究が扱う往信-返信文書間における文対応は異なった傾向を持っている。すなわち、新聞記事では同じ事象に対して複数の書き手が記事を作成したり、事象の経過により状況が異なったりすることで文書間や文書を跨いだ文間に対応が発生するのに対し、往信-返信文書ではコミュニケーションと言う目的を達成するために文対応が発生すると言う違いがある。また、本研究が対象としている往信-返信文書対における文対応では、4.2.3 節でも見た通り類似しない文同士にも文対応が存在することもあるため、対応する文同士が類似していることを前提にせず推定を行う必要がある。

新聞記事以外では、地方自治体間の条例を対象とした研究 (竹中・若尾, 2012)、料理レシピと対応するレビュー文書を対象とした研究 (Druck & Pang, 2012) がある。竹中・若尾 (2012) は地方自治体間で異なる条例を条文単位で比較する条文対応表を作成するために、条文間に対応を自動で推定する手法を提案している。また Druck & Pang (2012) は、レシピに対応するレビュー文書に含まれる作り方や材料に対する改善提案文の抽出を目的とし、その最終過程で提案文をレシピの手順と対応付ける手法を提案している。ただし、推定すべき対応が類似していることを前提としている (すなわち、竹中・若尾 (2012) の場合は同一の事柄に関する条例に対応付ける手法であり、Druck & Pang (2012) の場合はレシピ手順とレビュー文を対応付ける手法である) ため、これらの手法も対応する文の間に同じ単語や表現が出現していることを前提としている。

対話を対象とした研究には、Boyer et al. (2009) による対話における発話対応関係の分析がある。彼女らは、対話における隣接対 (adjacency pair) 構造を隠れマルコフモデル (Hidden Markov Model; HMM) を用いてモデル化している。ただし、彼女らの分析では、隣接対の場合は多くが位置的に隣接している可能性が高いことを前提としている^{*13}。これに対し、本研究の対象である文書対における文対応ではこう言った傾向を利用できないと言う違いがあるため、単純に彼女らの分析手法を本研究が対象としている文対応に適用することはできない。

^{*13} 隣接対の多くは位置的に隣接しているが、位置的に隣接していない場合もある。例えば、挿入連鎖 (隣接対の間に別の隣接対が挿入されるような構造) の場合は、位置的には離れた隣接対が観察される。

また、質問応答ウェブサイトの対話を対象とした研究として、Qu & Liu (2012) の質問応答ウェブサイトにおける文依存関係 (sentence dependency; 質問に対する回答、回答に対する解決報告など) を推定する研究がある。彼らは CRF による分類器を利用し、隠れマルコフモデルよりも高い性能で文依存関係を特定できたとしている。Qu & Liu (2012) と本研究では、異なる書き手 (発話者) による文間の関係を特定する点で類似しているが、以下の点で異なっている。まず、対象とする文書については、彼らの対象としているウェブサイトは対話に近い形で問題解決を図ると言う特徴を持っている一方で、本研究の対象は一对のやりとりでコミュニケーションが完結すると言う特徴を持つ。また、彼らが文依存関係の推定に予め推定した文種類を素性として投入しているのに対し、本研究はこれらを同時に推定する combine モデルを新たに提案している点でも異なる。

4.7 まとめと今後の課題

本論文では、対応する二文書間において文対応を自動で推定する課題を提案した。また、対話文書を対象とした従来手法を本課題に適用すると共に、文種類と文対応を推定するモデルを統合した新しいモデルを提案した。実際に文対応の推定性能について比較実験を行い、中再現率において統合モデルは他モデルよりも高い適合率であること、特に F 値最大点では最も高い F 値であることを確認した。

今後の課題として、以下の項目を考えている。

まず、本研究の対象は二文書の対に限っていたが、メールなどの場合では「返信の返信」のように三文書以上が関係する場合もある。ここで、往信文書を文書 A、文書 A に対する返信文書を文書 B、文書 B に対する再返信文書を文書 C とした場合、文書 B では文書 A に対する返信文に加え、文書 C への往信文が登場すると言う性質を持つ。例えば、文書 B における回答は文書 A と対応し、文書 B における質問は文書 C と対応する。また、文書 B における回答文に対して文書 C でフィードバックが行われている場合など、文書 B のある一文が文書 A, C 双方と関係を持つ場合もある。この場合、素朴な方法としては文書 A-文書 B 及び文書 B-文書 C のそれぞれで本研究での提案手法を繰り返し適用する方法がある。この際、文書 B の文種類をそれぞれの推定手順で返信文種類集合、往信文種類集合に切り替えて別々に推定すると言う方法もあるが、新たに往信文書・返信文書のいずれでもあるような文書のための文種類を新たに定義すると言う方法もある。更に、繰り返し推定するのではなく、文書 A, B, C の文種類・文対応を同時に推定するモデルに拡張すると言う方法を考えることもできる。加えて、本研究で扱う文対応の定義からは外れるものの、文書 B を経由せずに文書 A と文書 C で関連していると言った、三文書以上が関わることで初めて観察される関係もある。例えば、文書 A でした質問のいくつかが文書 B で答えられなかった場合に、文書 C で再度質問に触れる場合などがある。今後、三文書以上になることで新たに発生する事象についてはこのような関係も含めて分析を行い、三文書以上における文対応推定に最適な手法を検討したいと考えている。

加えて、本章の冒頭 4.1 節で紹介したような応用についても取り掛かりたいと考えている。今回対象となったデータセットであるレビュー文書・応答文書対についても様々な応用が考えられるため、応答文書の書き手の支援にとどまらず、ウェブサイトの利用者全体にとって有用なアプリケーションも実現したいと考えている。

第5章

おわりに

本研究では、EC サイトに関連するテキストを対象とした構想過程における文書作成支援として、以下の2つの研究課題を提案した。

1. 商品紹介文に相応しい評価視点の提示（第3章）
2. 対をなす二文書間における文対応関係の推定（第4章）

研究課題1では、商品紹介文に対する相応しい評価視点グループの提示を行う課題を提案し、評価視点表現の評価視点グループへの対応付け・評価視点グループのランキングへ分割する枠組みを提案した。評価実験を通じて提案した枠組みの有効性を実証し、結果の分析を通じて評価視点グループのランキングにおいては対数尤度比による手法が有効であり、評価視点グループ頻度のカウント手法については影響が小さいことを明らかにした。

研究課題2では、レビュー文書に対する応答文書など、二文書間における文対応関係の推定を行う課題を提案した。文対応の推定手法には、予め文種類を推定し、推定された文種類を文対応推定の素性として投入する手法と、文種類と文対応を推定するモデルを統合した新たなモデルを提案した。実際に文対応の推定性能について比較実験を行い、特に中再現率においては統合モデルが他モデルよりも高い適合率であることを示した。

以上に挙げた研究課題は、いずれも新たにEC サイトに関連したテキストの構想過程における文書作成支援を実現する課題であり、支援対象の拡大と言う方向により文書作成支援を発展させたと考える。加えて、文書作成支援に必要な情報源としてレビュー文書を用いており、レビュー文書の新たな利用可能性を示せたと考えている。

今後の課題を以下に挙げる。最初に第1章で、文書作成支援の発展の方向は「性能の向上」と「支援対象の拡大」であることを述べた。そこで、第一の課題としては、今回提案した文書作成支援における性能向上であり、具体的な指針は第3章、第4章の各章における考察で示した通りである。

第二の課題は、支援対象を更に拡大することである。第1章で、EC サイトに関連して販売者が作成するテキストを挙げたが、本研究では対象としなかったテキストもある。これらの中には、本研究と同様にレビュー文書を作成支援の情報源として活用できるものもあるため、新たな文書作成支援の可能性について検討したいと考えている。

付録 A

レビュー文書・応答文書における文分割手法

一般に、日本語文書は句点が文の区切れと対応している。しかし、ウェブ上の文書はくだけた表現やウェブサイト独自の記法が存在するため、必ずしも句点と文の句切れが対応しない。そのため、句点のみで文分割を行うと正しい分割結果が得られない場合がある。例えば、下図 A.1 は、あるレビュー文書を句点のみで分割した例である。図 A.1 に示すレビュー文書では、文の区切りとして記号「!」「★」やくだけた表現「(笑)」が使われているため、句点のみでは分割できない。

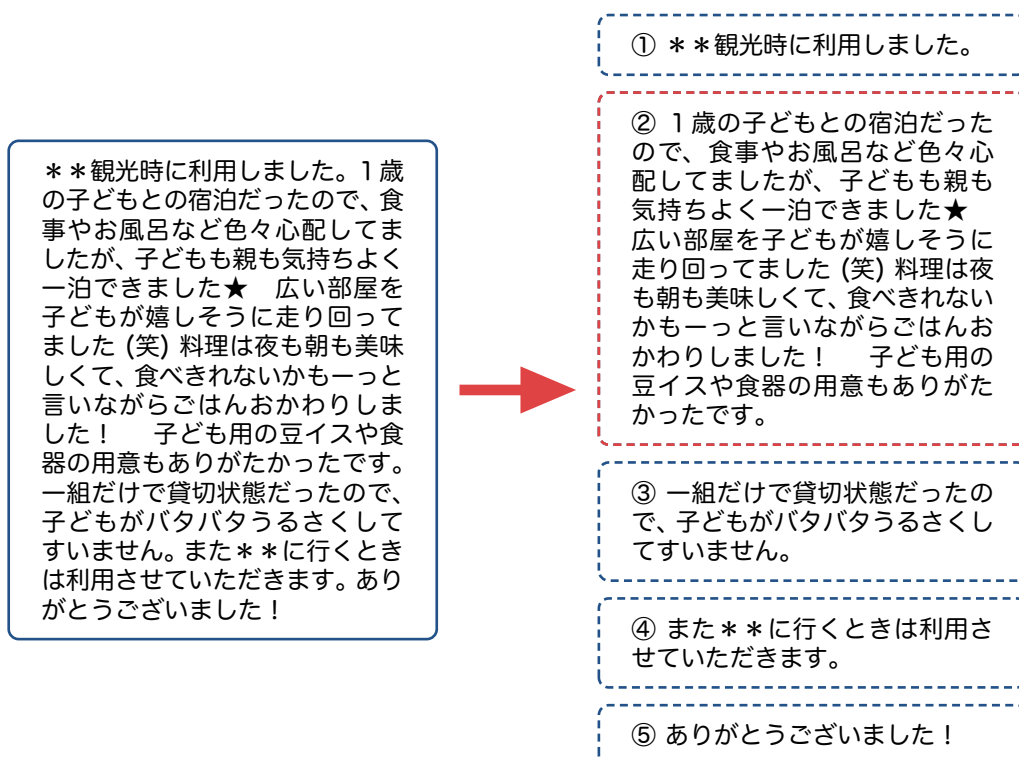


図 A.1: 句点のみで文分割を行った例

そこで、本研究で用いる文書、具体的には楽天トラベルのレビュー文書・応答文書に特化した文分割

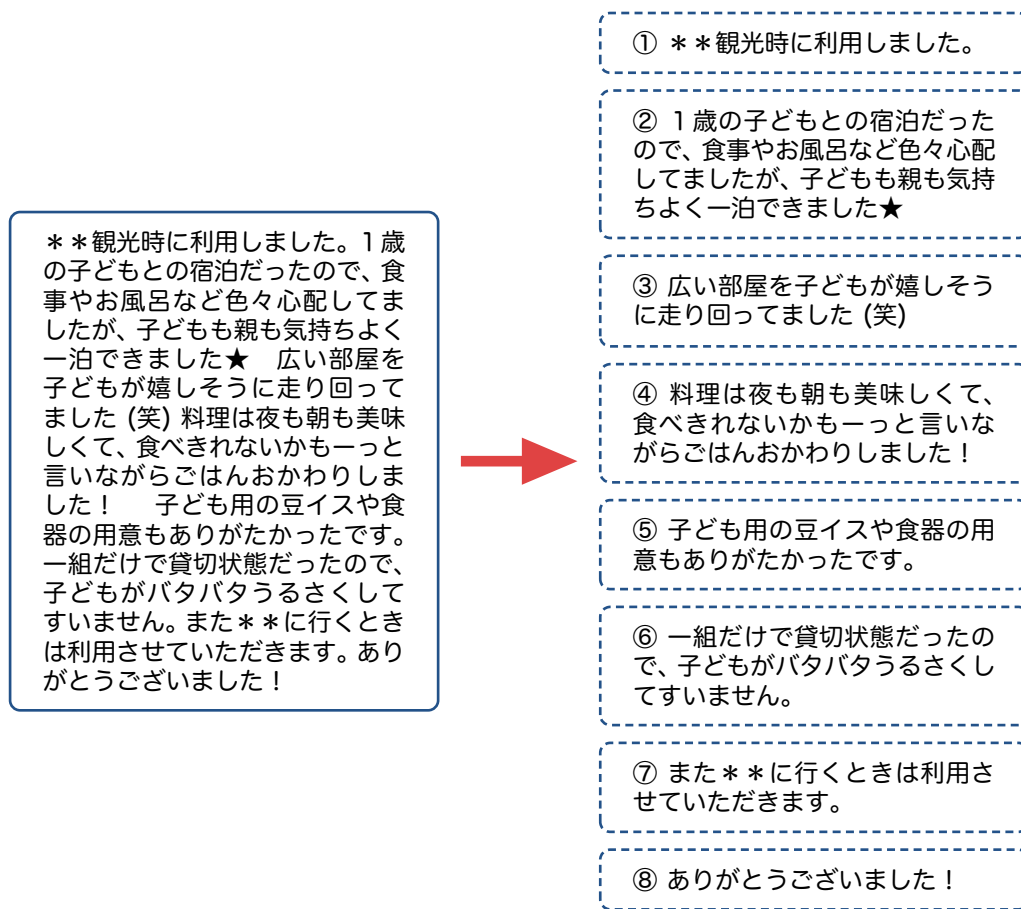


図 A.2: 本章で示す手法により文分割を行った例

ルールを作成した。予備実験の結果、単純なルールの組み合わせでも、楽天トラベルのようにくだけた表現が多用されていない文書であれば良好な分割結果が得られている。図 A.1 に示したレビュー文書を、提案する分割手法により分割した結果を図 A.2 に示す。図 A.2 より、正しい分割結果を得られていることが分かる。

以下、文分割に利用したルールについて述べる。例はいずれも実際の楽天トラベルの投稿から一部を引用したものであり、下線はその文字列が分割されないことを、波線部は分割点を表す。

レビュー文書・応答文書共通の分割ルール

基本分割文字列

次の各文字列で分割を行う。

?
!
。
.
♪
☆
★
*
※
(笑)
(泣)
(涙)

例外 (小数点)

. の前後が数値であった場合は分割を行わない。

例: 和室3.5畳 / 1.5リットル以上のペットボトル

例外 (No.N)

□ の前に文字列 No があり、後ろが数値であった場合は分割を行わない。

例: おすすめ度No.1 / ROOM No.411

例外 (句点-助詞)

MeCab (Kudo et al., 2004) による形態素解析の結果、基本分割文字列の後ろが助詞だった場合は分割を行わない。

例: 合宿免許? の若者さん達でしょうか / スタッフ? と話し込み

括弧の扱い (1)

括弧内に次の文字列があった場合は、括弧及び括弧内の文字列を一文とする。

□ □ (笑) □ (泣) □ (涙)

例: ~ (近日中には冷房に切り替わる予定です。) ~

1時間飲み放題 (カクテル各種!!) はお勧め

括弧の扱い (2)

括弧内に、**括弧の扱い (1)** の文字列ではない基本分割文字列が現れた場合は、二回以上登場した際に限り分割する。

例: ~ (セルフドリンクサービスはすごく良かったです! ~ 種類も豊富。) ~

レビュー文書のみでの分割ルール

例外 (宿泊プラン名)

【ご利用の宿泊プラン】 で強制的に分割し、以降は分割を行わない。

例: ~ 【ご利用の宿泊プラン】 ★オープニング特別★カップル&ファミリープラン

例外 (レーティング表記 1)

☆ 又は ★ の前が数値か、数値+つであった場合は分割を行わない。

例: 総合評価としては 5★ / 5つ☆ ホテル

例外 (レーティング表記 2)

☆ 又は ★ が連続して2つ以上並んでいる場合は分割を行わない。

例: それが無くても ★★★★★ です / とても寛げました。 (★★★★★)

応答文書のみでの分割ルール

署名区切り線

以下の文字が4文字以上連続で登場した場合、署名の区切りとみなして強制的に分割し、以降は分割を行わない。

— ~ = - * ☆ ★ + □ ■ ◇ ◆ ※ ∞

例: ===== ホテル***** フロント

参考文献

- [Asur & Huberman, 2010] Sitaram Asur, Bernardo A. Huberman. “Predicting the Future with Social Media.” in *Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT-2010)*, pp. 492–499, September, 2010.
- [Banko et al., 2000] Michele Banko, Vibhu O. Mittal, Michael J. Witbrock. “Headline Generation Based on Statistical Translation.” in *Proceedings of the 38th Annual Meeting on Association for Computational Linguistics (ACL-2000)*, pp. 318–325, July, 2000.
- [Bitner et al., 1990] Mary Jo Bitner, Bernard H. Booms, Mary Stanfield Tetreault. “The Service Encounter: Diagnosing Favorable and Unfavorable Incidents.” *Journal of Marketing*, Vol. 54, No. 1, pp. 71–84, January, 1990.
- [Bogdanova et al., 2012] Dasha Bogdanova, Paolo Rosso, Thamar Solorio. “On the Impact of Sentiment and Emotion Based Features in Detecting Online Sexual Predators.” in *Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis (WASSA-2012)*, pp. 110–118, July, 2012.
- [Boser et al., 1992] Bernhard E. Boser, Isabelle M. Guyon, Vladimir N. Vapnik. “A Training Algorithm for Optimal Margin Classifiers.” in *Proceedings of the 5th Annual Workshop on Computational Learning Theory (COLT-1992)*, pp. 144–152, July, 1992.
- [Boyer et al., 2009] Kristy Elizabeth Boyer, Robert Phillips, Eun Young Ha, Michael D. Wallis, Mladen A. Vouk, James C. Lester. “Modeling Dialogue Structure with Adjacency Pair Analysis and Hidden Markov Models.” in *Proceedings of the 2009 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL/HLT-2009)*, pp. 49–52, June, 2009.
- [Bringsjord & Ferrucci, 1999] Selmer Bringsjord, David Ferrucci. *Artificial Intelligence and Literary Creativity: Inside the Mind of Brutus, a Storytelling Machine*, Hillsdale, NJ, USA: L. Erlbaum Associates Inc. 1999.
- [Brody & Elhadad, 2010] Samuel Brody, Noemie Elhadad. “An unsupervised aspect-sentiment model for online reviews.” in *Proceedings of the 2010 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL/HLT-2010)*, pp. 804–812, June, 2010.
- [Bustamante & León, 1996] Flora Ramírez Bustamante, Fernando Sánchez León. “GramCheck: A grammar and style checker.” in *Proceedings of the 16th International Conference on Computational Linguistics (COLING-1996)*, pp. 175–181, August, 1996.
- [Byrd et al., 1994] Richard H. Byrd, Jorge Nocedal, Robert B. Schnabel. “Representations of quasi-Newton

- matrices and their use in limited memory methods.” *Mathematical Programming*, Vol. 63, No. 1-3, pp. 129–156, January, 1994.
- [Callaway & Lester, 2001] Charles B. Callaway, James C. Lester. “Narrative Prose Generation.” in *Proceedings of the 17th International Joint Conference on Artificial Intelligence (IJCAI-2001)*, pp. 1241–1250, August, 2001.
- [Carenini et al., 2005] Giuseppe Carenini, Raymond T. Ng, Ed Zwart. “Extracting knowledge from evaluative text.” in *Proceedings of the 3rd international conference on Knowledge capture (K-CAP-2005)*, pp. 11–18, October, 2005.
- [Carroll et al., 1998] John Carroll, Guido Minnen, Yvonne Canning, Siobhan Devlin, John Tait. “Practical Simplification of English Newspaper Text to Assist Aphasic Readers.” in *Proceedings of the 15th National Conference on Artificial Intelligence (AAAI-1998) Workshop on 10th Conference on Integrating Artificial Intelligence and Assistive Technology*, pp. 7–10, July, 1998.
- [Chen & Rosenfeld, 1999] Stanley F. Chen, Ronald Rosenfeld. “A gaussian prior for smoothing maximum entropy models.” Technical report, Carnegie Mellon University, 1999.
- [Chen et al., 2013a] Zhiyuan Chen, Arjun Mukherjee, Bing Liu, Meichun Hsu, Malu Castellanos, Riddhiman Ghosh. “Leveraging multi-domain prior knowledge in topic models.” in *Proceedings of the 23rd international joint conference on Artificial Intelligence (IJCAI-2013)*, pp. 2071–2077, August, 2013a.
- [Chen et al., 2013c] Zhiyuan Chen, Arjun Mukherjee, Bing Liu, Meichun Hsu, Malu Castellanos, Riddhiman Ghosh. “Exploiting Domain Knowledge in Aspect Extraction.” in *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP-2013)*, pp. 1655–1667, October, 2013c.
- [Chen et al., 2013b] Zhiyuan Chen, Arjun Mukherjee, Bing Liu, Meichun Hsu, Malu Castellanos, Riddhiman Ghosh. “Discovering coherent topics using general knowledge.” in *Proceedings of the 22nd ACM International Conference on Information and Knowledge Management (CIKM-2013)*, pp. 209–218, October, 2013b.
- [Chevalier & Mayzlin, 2006] Judith A. Chevalier, Dina Mayzlin. “The Effect of Word of Mouth on Sales: Online Book Reviews.” *Journal of Marketing Research*, Vol. 43, No. 3, pp. 345–354, August, 2006.
- [Coleman & Liau, 1975] Meri Coleman, Ta Lin Liau. “A computer readability formula designed for machine scoring.” *Journal of Applied Psychology*, Vol. 60, No. 2, pp. 283–284, April, 1975.
- [Cortes & Vapnik, 1995] Corinna Cortes, Vladimir Vapnik. “Support-Vector Networks.” *Machine Learning*, Vol. 20, No. 3, pp. 273–297, September, 1995.
- [Dale & Chall, 1948] Edgar Dale, Jeanne S. Chall. “A Formula for Predicting Readability.” *Educational Research Bulletin*, Vol. 27, No. 1, pp. 11–28, January, 1948.
- [Damerau, 1964] Fred J. Damerau. “A Technique for Computer Detection and Correction of Spelling Errors.” *Communications of the ACM*, Vol. 7, No. 3, pp. 171–176, March, 1964.
- [Darragh et al., 1990] John J. Darragh, Ian H. Witten, Mark L. James. “The Reactive Keyboard: a predictive typing aid.” *Computer*, Vol. 23, No. 11, pp. 41–49, November, 1990.
- [Dempster et al., 1977] A. P. Dempster, N. M. Laird, D. B. Rubin. “Maximum likelihood from incomplete data via the EM algorithm.” *Journal of the Royal Statistical Society: Series B*, Vol. 39, No. 1, pp. 1–38,

- January, 1977.
- [Druck & Pang, 2012] Gregory Druck, Bo Pang. “Spice it up?: mining refinements to online instructions from user generated content.” in *Proceedings of the 50th Annual Meeting of the Association of Computational Linguistics (ACL-2012)*, pp. 545–553, July, 2012.
- [Dunning, 1993] Ted Dunning. “Accurate Methods for the Statistics of Surprise and Coincidence.” *Computational Linguistics*, Vol. 19, No. 1, pp. 61–74, March, 1993.
- [Flesch, 1974] Rudolf Flesch. *The Art of Readable Writing*: Harper and Row, 1974.
- [Flower & Hayes, 1981] Linda Flower, John R. Hayes. “A Cognitive Process Theory of Writing.” *College Composition and Communication*, Vol. 32, No. 4, pp. 365–387, December, 1981.
- [Golding & Roth, 1999] Andrew R. Golding, Dan Roth. “A Winnow-Based Approach to Context-Sensitive Spelling Correction.” *Machine Learning*, Vol. 34, No. 1, pp. 107–130, February, 1999.
- [Golding & Schabes, 1996] Andrew R. Golding, Yves Schabes. “Combining Trigram-based and feature-based methods for context-sensitive spelling correction.” in *Proceedings of the 34th Annual Meeting on Association for Computational Linguistics (ACL-1996)*, pp. 71–78, June, 1996.
- [Green Jr. et al., 1961] Bert F. Green Jr., Alice K. Wolf, Carol Chomsky, Kenneth Laughery. “Baseball: An Automatic Question-answerer.” in *Proceedings of the Western Computing Conference, IRE-AIEE-ACM 1961 (Western)*, pp. 219–224, May, 1961.
- [Guo et al., 2009] Honglei Guo, Huijia Zhu, Zhili Guo, XiaoXun Zhang, Zhong Su. “Product feature categorization with multilevel latent semantic association.” in *Proceedings of the 18th ACM International Conference on Information and Knowledge Management (CIKM-2009)*, pp. 1087–1096, November, 2009.
- [Hatzivassiloglou & McKeown, 1997] Vasileios Hatzivassiloglou, Kathleen R. McKeown. “Predicting the Semantic Orientation of Adjectives.” in *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics (ACL/EACL-1997)*, pp. 174–181, July, 1997.
- [Hocutt et al., 2006] Mary Ann Hocutt, Michael R. Bowers, D. Todd Donovan. “The art of service recovery: fact or fiction?.” *Journal of Services Marketing*, Vol. 20, No. 3, pp. 199–207, April, 2006.
- [Holm, 1979] Sture Holm. “A Simple Sequentially Rejective Multiple Test Procedure.” *Scandinavian Journal of Statistics*, Vol. 6, No. 2, pp. 65–70, 1979.
- [Hu & Liu, 2004] Minqing Hu, Bing Liu. “Mining opinion features in customer reviews.” in *Proceedings of the 19th national conference on Artificial intelligence (AAAI-2004)*, pp. 755–760, July, 2004.
- [Hu et al., 2008] Nan Hu, Ling Liu, JieJennifer Zhang. “Do online reviews affect product sales? The role of reviewer characteristics and temporal effects.” *Information Technology and Management*, Vol. 9, No. 3, pp. 201–214, September, 2008.
- [Järvelin & Kekäläinen, 2002] Kalervo Järvelin, Jaana Kekäläinen. “Cumulated Gain-based Evaluation of IR Techniques.” *ACM Transactions on Information Systems*, Vol. 20, No. 4, pp. 422–446, September, 2002.
- [Jones & Martin, 1997] Michael P. Jones, James H. Martin. “Contextual Spelling Correction Using Latent Semantic Analysis.” in *Proceedings of the 5th Conference on Applied Natural Language Processing*

- (ANLP-1997), pp. 166–173, April, 1997.
- [Katz et al., 2004] Boris Katz, Matthew Bilotti, Sue Felshin, Aaron Fernandes, Wesley Hildebrandt, Roni Katzir, Jimmy Lin, Daniel Loreto, Gregory Marton, Federico Mora, Ozlem Uzuner. “Answering Multiple Questions on a Topic from Heterogeneous Resources.” in *Proceedings of the 13th Text Retrieval Conference (TREC-2004)*, November, 2004.
- [Katz et al., 2005] Boris Katz, Gary Borchardt, Sue Felshin. “Syntactic and Semantic Decomposition Strategies for Question Answering from Multiple Resources.” in *Proceedings of the 20th National Conference on Artificial Intelligence (AAAI-2005) Workshop on Inference for Textual Question Answering*, July, 2005.
- [Kau & Loh, 2006] Ah-Keng Kau, Elizabeth Wan-Yiun Loh. “The effects of service recovery on consumer satisfaction: a comparison between complainants and non-complainants.” *Journal of Services Marketing*, Vol. 20, No. 2, pp. 101–111, 2006.
- [Kleinberg, 1999] Jon M. Kleinberg. “Authoritative Sources in a Hyperlinked Environment.” *Journal of the ACM*, Vol. 46, No. 5, pp. 604–632, September, 1999.
- [Kschischang et al., 2001] Frank R. Kschischang, Brendan J. Frey, Hans-Andrea Loeliger. “Factor Graphs and the Sum-product Algorithm.” *IEEE Transactions on Information Theory*, Vol. 47, No. 2, pp. 498–519, February, 2001.
- [Kudo et al., 2004] Taku Kudo, Kaoru Yamamoto, Yuji Matsumoto. “Applying Conditional Random Fields to Japanese Morphological Analysis.” in *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP-2004)*, pp. 230–237, July, 2004.
- [Lafferty et al., 2001] John Lafferty, Andrew McCallum, Fernando Pereira. “Conditional random fields: Probabilistic models for segmenting and labeling sequence data.” in *Proceedings of the 18th International Conference on Machine Learning (ICML-2001)*, pp. 282–289, June, 2001.
- [Liu, 2012] Bing Liu. *Sentiment Analysis and Opinion Mining*: Morgan & Claypool Publishers, 2012.
- [Liu et al., 2007] Yang Liu, Xiangji Huang, Aijun An, Xiaohui Yu. “ARSA: A Sentiment-aware Model for Predicting Sales Performance Using Blogs.” in *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR-2007)*, pp. 607–614, July, 2007.
- [Lovelock & Wright, 1999] Christopher H. Lovelock, Lauren Wright. *Principles of Service Marketing and Management*: Prentice Hall, 1999.
- [Mangu & Brill, 1997] Lidia Mangu, Eric Brill. “Automatic Rule Acquisition for Spelling Correction.” in *Proceedings of the 14th International Conference on Machine Learning (ICML-1997)*, pp. 187–194, July, 1997.
- [Mann & Thompson, 1987] William C. Mann, Sandra A. Thompson. “Rhetorical Structure Theory: Description and Construction of Text Structures.” in Gerard Kempen ed. *Natural Language Generation*: Springer Netherlands, Chap. 7, pp. 85–95, 1987.
- [Mann et al., 1992] William C. Mann, Christian M.I.M. Matthiessen, Sandra A. Thompson. “Rhetorical Structure Theory and Text Analysis.” in William C. Mann, Sandra A. Thompson eds. *Discourse Description: Diverse linguistic analyses of a fund-raising text*: John Benjamins Publishing, pp. 39–78,

- 1992.
- [Manning & Schütze, 1999] Christopher D. Manning, Hinrich Schütze. *Foundations of Statistical Natural Language Processing*, Cambridge, MA, USA: MIT Press, 1999.
- [Marcu, 1997] Daniel Marcu. “The rhetorical parsing of natural language texts.” in *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics (ACL-1997)*, pp. 96–103, July, 1997.
- [Marcu, 1999] Daniel Marcu. “Discourse Trees Are Good Indicators of Importance in Text.” in Inderjeet Mani, Mark T. Maybury eds. *Advances in Automatic Text Summarization*: The MIT Press, Chap. 11, pp. 123–136, 1999.
- [Marcu et al., 2000] Daniel Marcu, Lynn Carlson, Maki Watanabe. “The automatic translation of discourse structures.” in *Proceedings of the 1st North American chapter of the Association for Computational Linguistics conference (NAACL-2000)*, pp. 9–17, April, 2000.
- [Mays et al., 1991] Eric Mays, Fred J. Damerau, Robert L. Mercer. “Context based spelling correction.” *Information Processing & Management*, Vol. 27, No. 5, pp. 517–522, 1991.
- [McCallum, 2002] Andrew Kachites McCallum. “MALLET: A Machine Learning for Language Toolkit.” <http://mallet.cs.umass.edu/>. 2002
- [McCallum & Li, 2003] Andrew McCallum, Wei Li. “Early Results for Named Entity Recognition with Conditional Random Fields, Feature Induction and Web-enhanced Lexicons.” in *Proceedings of the Conference on Computational Natural Language Learning (CoNLL-2003)*, pp. 188–191, June, 2003.
- [Meehan, 1976] James Richard Meehan. “The Metanovel: Writing Stories by Computer.” Ph.D. dissertation, Yale University, New Haven, CT, USA, 1976.
- [Michel & Meuter, 2008] Stefan Michel, Matthew L. Meuter. “The service recovery paradox: true but overrated?.” *International Journal of Service Industry Management*, Vol. 19, No. 4, pp. 441–457, August, 2008.
- [Mishne & Glance, 2006] Gilad Mishne, Natalie Glance. “Predicting Movie Sales from Blogger Sentiment.” in *Proceedings of the AAAI Spring Symposium: Computational Approaches to Analysing Weblogs*, pp. 155–158, March, 2006.
- [Mitton, 1996] Roger Mitton. *English spelling and the computer*: Harlow, Essex: Longman Group, 1996.
- [Mohammad & Yang, 2011] Saif M. Mohammad, Tony (Wenda) Yang. “Tracking Sentiment in Mail: How Genders Differ on Emotional Axes.” in *Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (WASSA-2011)*, pp. 70–79, June, 2011.
- [Neyman & Pearson, 1928] J. Neyman, E. S. Pearson. “On the Use and Interpretation of Certain Test Criteria for Purposes of Statistical Inference: Part I.” *Biometrika*, Vol. 20A, No. 1/2, pp. 175–240, July, 1928.
- [Nigam et al., 2000] Kamal Nigam, Andrew Kachites McCallum, Sebastian Thrun, Tom Mitchell. “Text Classification from Labeled and Unlabeled Documents using EM.” *Machine Learning*, Vol. 39, No. 2-3, pp. 103–134, May, 2000.
- [Oh & Sheng, 2011] Chong Oh, Olivia Sheng. “Investigating Predictive Power of Stock Micro Blog Sentiment in Forecasting Future Stock Price Directional Movement.” in *Proceedings of the 2011 International Conference on Information Systems (ICIS-2011)*, December, 2011.

- [Pang & Lee, 2005] Bo Pang, Lillian Lee. "Seeing Stars: Exploiting Class Relationships for Sentiment Categorization with Respect to Rating Scales." in *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics (ACL-2005)*, pp. 115–124, June, 2005.
- [Pang et al., 2002] Bo Pang, Lillian Lee, Shivakumar Vaithyanathan. "Thumbs Up?: Sentiment Classification Using Machine Learning Techniques." in *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP-2002)*, pp. 79–86, July, 2002.
- [Pedregosa et al., 2011] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, Édouard Duchesnay. "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research*, Vol. 12, pp. 2825–2830. <http://scikit-learn.org/>. October, 2011.
- [Peterson, 1980] James L. Peterson. "Computer Programs for Detecting and Correcting Spelling Errors." *Communications of the ACM*, Vol. 23, No. 12, pp. 676–687, December, 1980.
- [Pollock & Zamora, 1983] J. J. Pollock, A. Zamora. "Collection and characterization of spelling errors in scientific and scholarly text." *Journal of the American Society for Information Science*, Vol. 34, No. 1, pp. 51–58, January, 1983.
- [Popescu & Etzioni, 2005] Ana-Maria Popescu, Oren Etzioni. "Extracting Product Features and Opinions from Reviews." in *Proceedings of the 2005 Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP-2005)*, pp. 339–346, October, 2005.
- [Qu & Liu, 2012] Zhonghua Qu, Yang Liu. "Sentence Dependency Tagging in Online Question Answering Forums." in *Proceedings of the 50th Annual Meeting of the Association of Computational Linguistics (ACL-2012)*, pp. 554–562, July, 2012.
- [Quattoni et al., 2004] Ariadna Quattoni, Michael Collins, Trevor Darrell. "Conditional random fields for object recognition." in *Proceedings of the 18th Annual Conference on Neural Information Processing Systems (NIPS-2004)*, pp. 1097–1104, December, 2004.
- [Radev, 2000] Dragomir R. Radev. "A common theory of information fusion from multiple text sources step one." in *Proceedings of the 1st SIGdial Workshop on Discourse and Dialogue (SIGDIAL-2000)*, Vol. 10, pp. 74–83, October, 2000.
- [Radev et al., 2002] Dragomir Radev, Weiguo Fan, Hong Qi, Harris Wu, Amardeep Grewal. "Probabilistic Question Answering on the Web." in *Proceedings of the 11th International Conference on World Wide Web (WWW-2002)*, pp. 408–419, May, 2002.
- [Ratnaparkhi, 1997] Adwait Ratnaparkhi. "A simple introduction to maximum entropy models for natural language processing." Technical report, Institute for Research in Cognitive Science, 1997.
- [Razavi et al., 2014] Amir H. Razavi, Stan Matwin, Joseph De Koninck, Ray Reza Amini. "Dream sentiment analysis using second order soft co-occurrences (SOSCO) and time course representations." *Journal of Intelligent Information Systems*, Vol. 42, No. 3, pp. 393–413, June, 2014.
- [Resnik et al., 2005] Philip Resnik, Aaron Elkiss, Ellen Lau, Heather Taylor. "The Web in Theoretical Linguistics Research: Two Case Studies Using the Linguist's Search Engine." in *Proceedings of the 31st*

- Meeting of the Berkeley Linguistics Society*, pp. 265–276, February, 2005.
- [Ritter et al., 2011] Alan Ritter, Colin Cherry, William B. Dolan. “Data-driven response generation in social media.” in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing (EMNLP-2011)*, pp. 583–593, July, 2011.
- [Senter & Smith, 1967] R. J. Senter, E. A. Smith. “Automated Readability Index.” Technical report, Aerospace Medical Research Laboratories, Wright-Patterson Air Force Base, Ohio, 1967.
- [Sha & Pereira, 2003] Fei Sha, Fernando Pereira. “Shallow Parsing with Conditional Random Fields.” in *Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL/HLT-2003)*, pp. 134–141, May, 2003.
- [Shin et al., 2015] Andrew Shin, Ryohei Sasano, Hiroya Takamura, Manabu Okumura. “Context-Dependent Automatic Response Generation Using Statistical Machine Translation Techniques.” in *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL/HLT-2015)*, pp. 1345–1350, June, 2015.
- [Smith & Goodwin, 1971] Sidney L. Smith, Nancy C. Goodwin. “Alphabetic Data Entry via the Touch-Tone Pad: A Comment.” *Human Factors: The Journal of the Human Factors and Ergonomics Society*, Vol. 13, No. 2, pp. 189–190, April, 1971.
- [Sordoni et al., 2015] Alessandro Sordoni, Michel Galley, Michael Auli, Chris Brockett, Yangfeng Ji, Margaret Mitchell, Jian-Yun Nie, Jianfeng Gao, Bill Dolan. “A Neural Network Approach to Context-Sensitive Generation of Conversational Responses.” in *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL/HLT-2015)*, pp. 196–205, June, 2015.
- [Strong, 1925] Edward K. Jr. Strong. *The psychology of selling and advertising*: McGraw-Hill Book, Inc. 1925.
- [Sutton, 2006] Charles Sutton. “GRMM: GRaphical Models in Mallet.” <http://mallet.cs.umass.edu/grmm/>. 2006.
- [Tanaka-Ishii & Nakagawa, 2005] Kumiko Tanaka-Ishii, Hiroshi Nakagawa. “A Multilingual Usage Consultation Tool Based on Internet Searching: More Than a Search Engine, Less Than QA.” in *Proceedings of the 14th International Conference on World Wide Web (WWW-2005)*, pp. 363–371, May, 2005.
- [Thurmair, 1990] Gregor Thurmair. “Parsing for Grammar and Style Checking.” in *Proceedings of the 13th International Conference on Computational Linguistics (COLING-1990)*, pp. 365–370, August, 1990.
- [Titov & McDonald, 2008a] Ivan Titov, Ryan McDonald. “Modeling online reviews with multi-grain topic models.” in *Proceedings 17th international conference on World Wide Web (WWW-2008)*, pp. 111–120, April, 2008a.
- [Titov & McDonald, 2008b] Ivan Titov, Ryan McDonald. “A Joint Model of Text and Aspect Ratings for Sentiment Summarization.” in *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL/HLT-2008)*, pp. 308–316, June, 2008b.
- [Tsugawa et al., 2015] Sho Tsugawa, Yusuke Kikuchi, Fumio Kishino, Kosuke Nakajima, Yuichi Itoh, Hiroyuki Ohsaki. “Recognizing Depression from Twitter Activity.” in *Proceedings of the 33rd Annual*

- ACM Conference on Human Factors in Computing Systems (CHI-2015)*, pp. 3187–3196, April, 2015.
- [Tu et al., 2013] Mei Tu, Yu Zhou, Chengqing Zong. “A Novel Translation Framework Based on Rhetorical Structure Theory.” in *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL-2013)*, pp. 370–374, August, 2013.
- [Tumasjan et al., 2010] Andranik Tumasjan, Timm Sprenger, Philipp Sandner, Isabell Welp. “Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment.” in *Proceedings of the 4th International AAAI Conference on Weblogs and Social Media (ICWSM-2010)*, May, 2010.
- [Turney, 2002] P. D. Turney. “Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews.” in *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL-2002)*, pp. 417–424, July, 2002.
- [Vapnik & Lerner, 1963] V. Vapnik, A. Lerner. “Pattern recognition using generalized portrait method.” *Automation and Remote Control*, Vol. 24, pp. 774–780, January, 1963.
- [Wainwright et al., 2001] Martin Wainwright, Tommi Jaakkola, Alan Willsky. “Tree-based reparameterization for approximate inference on loopy graphs.” in *Proceedings of the 15th Annual Conference on Neural Information Processing Systems (NIPS-2001)*, pp. 1001–1008, December, 2001.
- [Wang et al., 2013] Yining Wang, Liwei Wang, Yuanzhi Li, Di He, Tie-Yan Liu, Wei Chen. “A Theoretical Analysis of NDCG Type Ranking Measures.” in *Proceedings of the 26th Annual Conference on Learning Theory (COLT-2013)*, June, 2013.
- [Wijaya, 2012] Bambang Sukma Wijaya. “The development of hierarchy of effects model in advertising.” *International Research Journal of Business Studies*, Vol. 5, No. 1, July, 2012.
- [Woods, 1973] William A. Woods. “Progress in Natural Language Understanding: An Application to Lunar Geology.” in *Proceedings of the American Federation of Information Processing Societies: 1973 National Computer Conference (AFIPS-1973)*, pp. 441–450, June, 1973.
- [Yamane & Hagiwara, 2015] Hiroaki Yamane, Masafumi Hagiwara. “Tag line generating system using knowledge extracted from statistical analyses.” *AI & SOCIETY*, Vol. 30, No. 1, pp. 57–67, February, 2015.
- [Yamane & Hagiwara, 2013] Hiroaki Yamane, Masafumi Hagiwara. “Tag Line Generating System Using Information on the Web.” *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 17, No. 2, pp. 185–193, March, 2013.
- [Yarowsky, 1994] David Yarowsky. “Decision Lists for Lexical Ambiguity Resolution: Application to Accent Restoration in Spanish and French.” in *Proceedings of the 32nd Annual Meeting on Association for Computational Linguistics (ACL-1994)*, pp. 88–95, June, 1994.
- [Yu & Hatzivassiloglou, 2003] Hong Yu, Vasileios Hatzivassiloglou. “Towards Answering Opinion Questions: Separating Facts from Opinions and Identifying the Polarity of Opinion Sentences.” in *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing (EMNLP-2003)*, pp. 129–136, July, 2003.
- [Yu et al., 2011] Jianxing Yu, Zheng-Jun Zha, Meng Wang, Tat-Seng Chua. “Aspect ranking: identifying important product aspects from online consumer reviews.” in *Proceedings of the 49th Annual Meeting*

- of the Association for Computational Linguistics: Human Language Technologies (ACL/HLT-2011)*, pp. 1496–1505, June, 2011.
- [Zhai et al., 2011] Zhongwu Zhai, Bing Liu, Hua Xu, Peifa Jia. “Clustering Product Features for Opinion Mining.” in *Proceedings of the 4th ACM International Conference on Web Search and Data Mining (WSDM-2011)*, pp. 347–354, February, 2011.
- [Zhai et al., 2010] Zhongwu Zhai, Bing Liu, Hua Xu, Peifa Jia. “Grouping product features using semi-supervised learning with soft-constraints.” in *Proceedings of the 23rd International Conference on Computational Linguistics (COLING-2010)*, pp. 1272–1280, August, 2010.
- [Zhang & Liu, 2014] Lei Zhang, Bing Liu. “Aspect and Entity Extraction for Opinion Mining.” in Wesley W. Chu ed. *Data Mining and Knowledge Discovery for Big Data: Methodologies, Challenges, and Opportunities*: Springer, pp. 1–40, 2014.
- [Zhang et al., 2010] Lei Zhang, Bing Liu, Suk Hwan Lim, Eamonn O’Brien-Strain. “Extracting and ranking product features in opinion documents.” in *Proceedings of the 23rd International Conference on Computational Linguistics (COLING-2010)*, pp. 1462–1470, August, 2010.
- [Zhu et al., 2005] Jun Zhu, Zaiqing Nie, Ji-Rong Wen, Bo Zhang, Wei-Ying Ma. “2D Conditional Random Fields for Web information extraction.” in *Proceedings of the 22nd international conference on Machine learning (ICML-2005)*, pp. 1044–1051, August, 2005.
- [浅原, 2007] 浅原正幸. 自然言語処理と系列ラベリング技術. オペレーションズ・リサーチ: 経営の科学, Vol. 52, No. 11, pp.689–694, 2007年11月.
- [荒木他, 2002] 荒木哲郎, 池原悟, 佐藤政伸, 榮代正男. マルコフ連鎖モデルを用いた日本語文の置換型, 挿入型及び脱落型誤りの検出・訂正法の改善. 電子情報通信学会論文誌, Vol. J85-DII, No. 1, pp.66–78, 2002年1月.
- [アンヴィコミュニケーションズ, 2015] アンヴィコミュニケーションズ. AISCEAS の法則とは, 2015年, <http://www.amviy.jp/aisceas/> (アクセス日: 2015年12月10日).
- [石岡・亀田, 2003] 石岡恒憲, 亀田雅之. コンピュータによる小論文の自動採点システム Jess の試作. 計算機統計学, Vol. 16, No. 1, pp.3–19, 2003年12月.
- [石塚, 2005] 石塚隆男. エンティティの自動抽出による英文記事の構造の可視化. 人工知能学会 2005年全国大会, 2005年6月.
- [乾他, 2013] 乾孝司, 板谷悠人, 山本幹雄, 新里圭司, 平手勇宇, 山田薫. 意見集約における相対的特徴を考慮した評価視点の構造化. 自然言語処理, Vol. 20, No. 1, pp.3–25, 2013年6月.
- [乾・奥村, 2006] 乾孝司, 奥村学. テキストを対象とした評価情報の分析に関する研究動向. 自然言語処理, Vol. 13, No. 3, pp.201–241, 2006年7月.
- [内山他, 2004] 内山将夫, 中條清美, 山本英子, 井佐原均. 英語教育のための分野特徴単語の選定尺度の比較. 自然言語処理, Vol. 11, No. 3, pp.165–197, 2004年7月.
- [大沢他, 2010a] 大沢裕子, 郷亜里沙, 安田励子. インターネット上のクチコミにおける苦情への返答—サイト閲覧者の視点から—. 待遇コミュニケーション研究, No. 7, pp.1–16, 2010a年1月.
- [大沢他, 2010b] 大沢裕子, 郷亜里沙, 安田励子. Web サイトにおけるクチコミの苦情と返答—「宿泊予約サイト」を対象に—. 言語処理学会第16回年次大会, pp.322–325, 2010b年3月.

- [小方, 2007] 小方孝. プロップから物語内容の修辞学へ—解体と再構成の修辞を中心として—. 認知科学, Vol. 14, No. 4, pp.532–558, 2007年12月.
- [奥野・萩原, 2009] 奥野陽, 萩原将文. インターネットを用いた日本語入力システム. 情報処理学会研究報告, 自然言語処理研究会報告, 2009-NL-190, Vol. 2009, No. 36, pp.1–6, 2009年3月.
- [影浦, 1997] 影浦峯. 文字単位の bigram 尺度に基づく複合漢字列の単位切り手法. 言語処理学会第3回年次大会, pp.477–480, 1997年3月.
- [笠原他, 1994] 笠原要, 松澤和光, 石川勉, 河岡司. 観点に基づく概念間の類似性判別. 情報処理学会論文誌, Vol. 35, No. 3, pp.505–509, 1994年3月.
- [鴨崎他, 2008] 鴨崎真直, 大野陽介, Ruck Thawonmas. 物語生成システムにおける動作への感情を考慮するキャラクターの配役とその効果. ゲーム学会論文誌, Vol. 2, No. 1, pp.36–41, 2008年.
- [工藤他, 2004] 工藤拓, 山本薫, 松本裕治. Conditional Random Fields を用いた日本語形態素解析. 情報処理学会研究報告, 自然言語処理研究会報告, 1998-NL-161, Vol. 2004, No. 47, pp.89–96, 2004年5月.
- [工藤他, 2011] 工藤拓, 小松弘幸, 花岡俊行, 向井淳, 田畑悠介. 統計的かな漢字変換システム Mozc. 言語処理学会第17回年次大会, pp.948–951, 2011年3月.
- [経済産業省, 2014] 経済産業省. 平成26年度電子商取引に関する市場調査, 2014年, http://www.meti.go.jp/policy/it_policy/statistics/outlook/ie_outlook.html (アクセス日: 2015年12月10日).
- [現代用語の基礎知識, 2015] JapanKnowledge Lib 現代用語の基礎知識. AISAS(アイサス), 2015年, <http://japanknowledge.com/lib/display/?lid=5002013600800> (アクセス日: 2015年12月10日).
- [小林・小木曾, 2013] 小林雄一郎, 小木曾智信. 中古和文における個人文体とジャンル文体: 多変量解析による歴史的資料の文体研究. 国立国語研究所論集, Vol. 6, pp.29–43, 2013年.
- [小松他, 2003] 小松弘幸, 高林哲, 増井俊之. 動的略語展開を利用した文脈をとらえた予測入力. 情報処理学会論文誌, Vol. 44, No. 11, pp.2538–2546, 2003年11月.
- [齋藤他, 2007] 齋藤邦子, 鈴木潤, 今村賢治. CRF を用いたブログからの固有表現抽出. 言語処理学会第13回年次大会, pp.107–110, 2007年3月.
- [佐久間・小方, 2005] 佐久間友子, 小方孝. プロップの物語内容論を利用したストーリー生成支援システムとその考察. 人工知能学会2005年全国大会, 2005年6月.
- [佐久間・小方, 2006] 佐久間友子, 小方孝. 行程規則を用いた複数のストーリーの合成—ストーリー自動生成機能を持つストーリー生成支援システムへの一アプローチ—. 人工知能学会2006年全国大会, 2006年6月.
- [下園他, 1994] 下園幸一, 菅沼明, 牛島和夫. 日本語文章推敲支援ツール『推敲』における助詞「が」の抽出について. 情報処理学会論文誌, Vol. 35, No. 8, pp.1652–1660, 1994年8月.
- [新納, 1999] 新納浩幸. 平仮名 N-gram による平仮名列の誤り検出とその修正. 情報処理学会論文誌, Vol. 40, No. 6, pp.2690–2698, 1999年6月.
- [菅沼他, 1990] 菅沼明, 倉田昌典, 牛島和夫. 日本語文章推敲支援ツール『推敲』における否定表現の抽出法. 情報処理学会論文誌, Vol. 31, No. 6, pp.792–800, 1990年6月.

- [菅沼・牛島, 1991] 菅沼明, 牛島和夫. 日本語文章推敲支援ツール『推敲』におけるとりたて詞「は」の抽出法とその評価. 情報処理学会論文誌, Vol. 32, No. 11, pp.1392-1400, 1991年11月.
- [鈴木・武田, 1989] 鈴木恵美子, 武田浩一. 日本語文書校正支援システムの設計と評価. 情報処理学会論文誌, Vol. 30, No. 11, pp.1402-1412, 1989年11月.
- [高橋・牛島, 1991] 高橋善文, 牛島和夫. 計算機マニュアルの分かりやすさの定量的評価方法. 情報処理学会論文誌, Vol. 32, No. 4, pp.460-469, 1991年4月.
- [竹中・若尾, 2012] 竹中要一, 若尾岳志. 地方自治体の例規比較に用いる条文対応表の作成支援. 自然言語処理, Vol. 19, No. 3, pp.193-212, 2012年9月.
- [建石他, 1988] 建石由佳, 小野芳彦, 山田尚勇. 日本文の読みやすさの評価式. 情報処理学会研究報告, ヒューマンコンピュータインタラクション研究会報告, 1988-HI-018, Vol. 1988, No. 25, pp.1-8, 1988年5月.
- [田村・和田, 1998] 田村直良, 和田啓二. セグメントの分割と統合による文章の構造解析. 自然言語処理, Vol. 5, No. 1, pp.59-78, 1998年1月.
- [恒川他, 1993] 恒川昌昭, 高尾直弥, 川原千尋, 辻敦生, 對馬常人, 杉村領一. 文書作成支援システム. 情報処理学会研究報告, 自然言語処理研究会報告, 1993-NL-95, Vol. 1993, No. 41, pp.125-132, 1993年5月.
- [中野・鬼沢, 2008] 中野俊亮, 鬼沢武久. ユーザ対話による意外性を持つキャッチフレーズ作成支援. 情報処理学会第70回全国大会, pp.201-202, 2008年3月.
- [長島, 2009] 長島広太. サービス商品の選択・評価におけるクチコミの機能-ホテルサービスの場合-. 経営論集, Vol. 74, pp.91-108, 2009年11月.
- [難波他, 2005] 難波英嗣, 国政美伸, 福島志穂, 相沢輝昭, 奥村学. 文書横断文間関係を考慮した動向情報の抽出と可視化. 情報処理学会研究報告, 自然言語処理研究会報告, 2005-NL-168, Vol. 2005, No. 73, pp.67-74, 2005年7月.
- [日本経済新聞, 2015] 日本経済新聞. アマゾン出店者17.8万、仮想モール、6月末時点、4～6月取扱高は1.4倍に, 2015年8月6日朝刊, 15ページ.
- [納富・内山, 1991] 納富一宏, 内山明彦. 自然言語処理を応用したマニュアル作成支援システム-マニュアル推敲支援に関して-. 情報処理学会研究報告, 自然言語処理研究会報告, 1991-NL-85, Vol. 1991, No. 80, pp.89-96, 1991年9月.
- [長谷川他, 2013] 長谷川貴之, 鍛冶伸裕, 吉永直樹, 豊田正史. 聞き手の感情を喚起する発話応答生成. 言語処理学会第19回年次大会, pp.150-153, 2013年3月.
- [幅野・浦谷, 2013] 幅野裕貴, 浦谷則好. 映画のキャッチコピー作成支援手法. 言語処理学会第19回年次大会, 2013年3月.
- [幅野・浦谷, 2014] 幅野裕貴, 浦谷則好. あらすじとレビューを用いた映画のキャッチコピー作成支援手法の提案. 言語処理学会第20回年次大会, 2014年3月.
- [林・菊井, 1991] 林良彦, 菊井玄一郎. 日本文推敲支援システムにおける書換え支援機能の実現方式. 情報処理学会論文誌, Vol. 32, No. 8, pp.962-970, 1991年.
- [兵藤他, 1996] 兵藤安昭, 河田実成, 応江黔, 池田尚志. 構文付きコーパスの作成と類似用例検索システムへの応用. 自然言語処理, Vol. 3, No. 2, pp.73-88, 1996年4月.

- [平尾他, 2013] 平尾努, 西野正彬, 安田宜仁, 永田昌明. 談話構造に基づく単一文書要約. 言語処理学会第 19 回年次大会, pp.492–495, 2013 年 3 月.
- [比留間他, 1999] 比留間正樹, 山下卓規, 奈良雅雄, 田村直良. 文章の構造化による修辭情報を利用した自動抄録と文章要約. 自然言語処理, Vol. 6, No. 6, pp.113–129, 1999 年 7 月.
- [藤村他, 2005] 藤村元彦, 新納浩幸, 佐々木稔. RSS フィード作成のためのニュース記事タイトルの抽出手法. 言語処理学会第 11 回年次大会, 2005 年 3 月.
- [松平・萩原, 2005] 松平智史, 萩原将文. 対話型遺伝的プログラミングと電子化辞書を用いたキャッチコピー作成支援システム. 電気学会論文誌 C, Vol. 125, No. 4, pp.616–622, 2005 年 4 月.
- [松平・萩原, 2004] 松平智史, 萩原将文. 電子化辞書と遺伝的プログラミングを用いたキャッチコピー作成支援システム. 電気学会論文誌 C, Vol. 124, No. 1, pp.164–169, 2004 年 1 月.
- [松本他, 2009] 松本章代, 山田未央佳, 山田翔, 鈴木雅人. 理工系学生を対象とした技術文書作成支援システム. 情報処理学会研究報告, コンピュータと教育研究会報告, 2009-CE-98, Vol. 2009, No. 15, pp.91–96, 2009 年 2 月.
- [三品他, 2004] 三品拓也, 貞光九月, 山本幹雄. 確率的 LSA を用いた日本語同音異義語誤りの検出・訂正. 情報処理学会論文誌, Vol. 45, No. 9, pp.2168–2176, 2004 年 9 月.
- [南野・奥村, 2005] 南野朋之, 奥村学. RSS 自動生成のためのタイトル生成. 言語処理学会第 11 回年次大会, 2005 年 3 月.
- [南他, 2011] 南和江, 藤井康寿, 土屋雅稔, 中川聖一. 大規模コーパスを用いた固有表現抽出手法の検討. 言語処理学会第 17 回年次大会, pp.328–331, 2011 年 3 月.
- [宮部他, 2006] 宮部泰成, 高村大也, 奥村学. 文書横断文間関係の特定. 言語処理学会第 12 回年次大会, pp.496–499, 2006 年 3 月.
- [宮部他, 2005] 宮部泰成, 高村大也, 奥村学. 異なる文書中の文間関係の特定. 情報処理学会研究報告, 自然言語処理研究会, 2005-NL-168, Vol. 2005, pp.97–104, 2005 年 7 月.
- [山崎, 1998] 山崎邦子. 聴覚障害者用字幕生成のための言い換えによるニュース文要約. 言語処理学会第 4 回年次大会, pp.646–649, 1998 年 3 月.
- [吉崎他, 2010] 吉崎智則, 宮村幸祐, 金田重郎. マルチエージェント・シミュレーションを用いたナラティブ生成システムの提案. 情報処理学会第 72 回全国大会, pp.665–666, 2010 年 3 月.
- [若尾他, 1997] 若尾孝博, 江原暉将, 白井克彦. テレビニュース番組の字幕に見られる要約の手法. 情報処理学会研究報告, 自然言語処理研究会報告, 1997-NL-122, Vol. 1997, No. 109, pp.83–89, 1997 年 11 月.
- [脇田・奥村, 1993] 脇田早紀子, 奥村薫. 日本語校正支援システム FleCS – ミスタイプ検出について. 情報処理学会研究報告, 自然言語処理研究会報告, 1993-NL-97, Vol. 1993, No. 79, pp.135–142, 1993 年 9 月.

謝辞

山本幹雄教授には、本研究を含め、学類時代からの長きにわたって丁寧かつ多大なる御指導を賜りました。研究面においては、様々な観点からのアドバイスにより、新たな可能性を様々示して頂きました。また、様々な面白い研究アイデアを思いつくだけでなく、アイデアを惜しげも無く私を含む学生に与えて下さいました。加えて、学生生活面においても、ティーチングアシスタントやリサーチアシスタントなどのお世話に加えて、各種学生支援においても、常に私にとって良い結果を得られるよう多忙な中で東奔西走して頂きました。以上のような話を何とはなしに知人にして「本当にいい先生だね」と言う言葉を聞いて、山本教授の御指導に与ることができて本当に良かったと改めて感じました。謹んで、深甚なる感謝の意を申し上げます。

乾孝司准教授には、本研究を進めるに当たって常日頃から御指導・御助言を賜りました。乾准教授には、私が何度も研究に行き詰まった時でも、多数の具体的なアドバイスにより常に道を示して頂きました。かの个性化的な評判分析グループのメンバーを以ってして、異口同音に「乾先生には足を向けて寝られない」と言わしめる程の丁寧な御指導を各人に向ける多忙な中にありながら、私にも懇切丁寧な御指導を頂きました乾准教授に厚く御礼申し上げます。

狩野均教授、福井和広教授、関洋平准教授の諸先生方からは、本学位論文の審査を通じて、数々の貴重な御助言を頂きました。先生方の的確かつ鋭い御指摘のお陰で、本論文は大いに改善されました。深く感謝致します。

楽天技術研究所ニューヨークでのインターンシップでは、楽天技術研究所の皆様にお世話になりました。特に Zofia Stankiewicz さんはメンターとして付いて頂き、インターン期間中のみならず、帰国してからも親身に研究のご相談に応じて頂きました。また、関根聡准教授にはインターンシップ期間中、研究のことにとどまらず、不慣れな海外で生活するに当たって様々なお世話を頂きました。実家を長期離れたのが初めて、さらに日本を離れたのも初めての出来事が3ヶ月間の海外インターンシップであったにも関わらず、無事にインターンシップを終えることができたのはお二人のお陰です。心より御礼申し上げます。

津川翔助教には、ゼミの場を中心としまして様々な御指導を頂きました。専門外でありながらも示唆に富み、そして親しみが湧くアドバイスは、研究を進めるにあたって大変励みになりました。深く感謝の意を表します。

本研究の実験にあたり、楽天データ公開において公開された楽天トラベル「お客さまの声・クチコミ」データを使用させて頂きました。貴重なデータを公開頂きました楽天株式会社に感謝を申し上げます。

長期に渡る研究生生活では、知能情報研究室の先輩・同輩・後輩の皆様方に、研究はもちろんのこと多方面においてお世話になりました。誠にありがとうございました。

また、博士後期課程への進学に一切の干渉を行わないばかりか、博士後期課程修了までの長期に渡って実家に住まうことを許して頂いた両親に感謝致します。

最後に、博士後期課程の在学期間を含む長年にわたり、常住不断の精神的支えを下さった坂元なつみさんに感謝申し上げます。私自身の博士課程進学と遠距離交際が重なる中で大変な不安を感じていたであろうにも関わらず、気丈かつ辛抱強い支えと、多大なる愛情を注いで頂きました。本当にありがとう。

公表済み論文一覧

学術雑誌（査読有り）

1. 角田孝昭, 乾孝司, 山本幹雄. 対をなす二文書間における文対応関係の推定. 自然言語処理. 22 (1), 2015 年 3 月. 研究課題 2 (第 4 章)

国際会議（査読有り）

1. Takaaki Tsunoda, Takashi Inui, Satoshi Sekine. Utilizing review analysis to suggest product advertisement improvements. 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA-2015), Lisboa, Portugal, September, 2015. 研究課題 1 (第 3 章)

国内学会

1. 角田孝昭, 吉田光男, 津川翔, 山本幹雄. 状態空間モデルを用いた検索トレンドとページビューからの自動車販売台数の予測. 人工知能学会 2015 年度 (第 29 回) 全国大会 (JSAI-2015), 3C3-3, 公立ほこだて未来大学, 2015 年 6 月.
2. 角田孝昭, Zofia Stankiewicz, 関根聡. 商品紹介文改善のための紹介文およびレビュー集合における評価視点の整理. 言語処理学会第 20 回年次大会 (NLP-2014), P8-1, pp. 876-879, 北海道大学, 2014 年 3 月.
3. 角田孝昭, 乾孝司, 山本幹雄. 対をなす二文書間における文対応推定および応答文生成への応用. 言語処理学会第 19 回年次大会 (NLP-2013), B3-3, pp. 198-201, 名古屋大学, 2013 年 3 月.
4. 角田孝昭, 乾孝司, 山本幹雄. 局所的及び大域的文脈を併用した日本語同音異義語誤りの訂正. 言語処理学会第 17 回年次大会 (NLP-2011), P3-14, pp. 635-638, 豊橋科学技術大学, 2011 年 3 月.

シンポジウム他

1. Takaaki Tsunoda. Utilizing review analysis to suggest product blurb improvement. ICT International Exchange Workshop 2014, University of Tsukuba, November, 2014.
2. 角田孝昭. 対象 Web サービスに特化した日本語スペルチェッカーの開発. 「大規模情報コンテンツ時代の高度 ICT 専門職業人育成」事業 第 1 回シンポジウム, 筑波大学, 2011 年 3 月.

3. 角田孝昭, 澤田健都, 吉田光男. 対象サービスに特化した日本語 Web スペルチェッカーの開発.
第3回楽天研究開発シンポジウム, 東京都, 楽天タワー2号館, 2010年12月.