

エンティティ間の類似関係取得のための Wikipedia 事象モデル構築手法に関する考察

内藤 稔[†] 浅野 泰仁[†] 吉川 正俊[†]

[†] 京都大学大学院情報学研究科社会情報学専攻 〒606-8501 京都府京都市左京区吉田本町

E-mail: [†]naito@db.soc.i.kyoto-u.ac.jp, ^{††}{asano,yoshikawa}@i.kyoto-u.ac.jp

あらまし エンティティ間の類似関係の取得は、情報の集約において重要な要素である。その取得において、現在は語知識や共起といった手法を用いているものが多いが、文章の流れや事象を考慮したものは少ない。そのため、直接的な単語の類似や共起の無い、人物の行動（例：経済危機に対して政策を打ち出し、成功した）による類似関係等が取得しづらい問題がある。そこで、我々は文書から述語項を基本とした最小事象を抽出し、その間の関係をグラフとして捉える事象モデルを考案した。これにより、事象の流れを考慮した、事象及びその事象を持つエンティティの比較が可能となり、これまでは困難であった関係の発見が可能になると考える。

キーワード エンティティ間関係, 自然言語処理

1. はじめに

近年、情報の集約という作業は重要性を増している。これは、ユーザが触れる情報の爆発的増加により、膨大な情報源から単純に情報を列挙するだけでは、ユーザの必要な情報の取得が困難になっているためである。そこで、ユーザは収集した情報を集約する必要がある。

このような問題は、人手によって集約された情報の出現により解決されつつある。そのような例として Wikipedia 等が挙げられる。しかし、人手による集約には限界があり、情報の不足や、間違いが混入する可能性も少なくない。また、集約の方法はエンティティ毎に纏めるものが多い。事実 Wikipedia 上では、エンティティ毎に記事を作成し、そのエンティティに関する情報はそのページ内に集約し、完結させている。エンティティ間は、記事中のアンカーリンクといったもので結ばれているが、その内容などを参照するものではない。そのため、複数のエンティティに跨った情報の集約や、エンティティの比較といったタスクにはまだ問題がある。

一方、エンティティ間の類似関係を取得する研究も多くある。例えば語知識データベース等を利用したものや、Wikipedia のアンカーリンクや年表情報を利用したものがある。これらは主に語ベースの関係取得の方法であり、既知の関係であったり、共起のある関係であったりする事が多く、事象や文章の流れを考慮した複雑な関係の取得は難しい。一例として、ナポレオン・ボナパルトとガイウス・ユリウス・カエサルは比較される事も多いが、Wikipedia の記事中で、相互に相手の名前を含まない。また、両者は生まれた時代背景も場所も違うため、機械的な単純な比較は難しい。しかし、人の目で見れば、例えば目的のために婚約者と分かれて結婚していたり、若い頃に国外脱出という逃避行を行っていたりと、共通するような行動も多い事が分かる。また別の例として、享保の大飢饉を受けて被害を

受けた藩政を立て直した伊達村候と、世界恐慌による経済不況をニューディール政策により建て直したフランクリン・ルーズベルトの両者も、その行動において類似した関係があると言える。このような例で重要なのは、以下のようにそれぞれ類似する要素が、同じ流れで展開されているという点にある。

- 経済危機によって大きな痛手を受けた。
- 一定の政策を行った
- 財政を立て直した

本稿では、このような要素を最小の事象として定義し、最小事象をノード、その間の関係をエッジとするグラフモデルを、エンティティが持つ事象モデルとして提案する。この部分グラフが全て事象であり、その比較を行う事で、記事中の文章の流れを考慮した事象の比較が行える。この手法によって、今まで発見が困難であった上で挙げたような類似関係が、その理由となる個所と共に発見できるようになると考える。

このような手法の実現のためには、以下のような問題に対して取り組む必要がある。

- 記事中から最小事象を抽出する
- 抽出した最小事象間の関係を取得する
- 取得した関係に従ってグラフを構築する
- 事象グラフ同士の比較を行う

このような課題を、本稿では Wikipedia 上の人物記事に特化した手法を提案する。Wikipedia、及び人物記事を用いるのは、以下に挙げるような特徴が、本手法にとって有用だと考えるためである。

- アンカーリンクによる固有表現の正規化が容易
- カテゴリ等による、記事の属性の同定が容易
- 記事は客観的な情報が基本で、主観的な文章が少ない
- 1記事1エンティティの対応であるため、エンティティの情報が1記事中で完結している
- 特に歴史的人物においては、経歴など事実の記述が多い

- 後述する事象性名詞が含まれやすい

このような特徴を利用して、Wikipedia の人物記事中から事象モデルを抽出する。

本稿の構成は以下の通りである。2 節では関連研究について述べ、本研究が取りくむ問題について明らかにする。3 節では本稿で扱う事象に対する定義、及び定義した最小事象の抽出方法について述べる。4 節では抽出した最小事象間の関係の取得方法及びグラフ化の手法について述べる。5 節では、本研究のこれからの課題について述べる。

2. 関連研究

松吉 [1] らは、大西らの動詞語釈文構造化データ・竹内らの動詞項構造シソーラスを用いて事象間関係知識データベースを作成し、Web 上の言論間にある関係をユーザに提示する言論マップの提案を行っている。この研究では、Web 上の言論に対する対立する言論や根拠となる言論等を提示することで、ユーザの情報選択・集約を支援するものである。

島田 [2] らは、Wikipedia Sentential Event(WSE) モデルを提案し、Wikipedia における同一事象検出を行う研究を行っている。この研究では、Wikipedia 記事の 1 文を 1 事象として扱い、記事中から事象情報データベースに事象モデルを格納している。この事象情報データベースから、同一事象と思われる事象を抽出し、情報の不整合等があればそれをユーザに提示し、ユーザの情報収集を助けるものである。

後藤 [3] らは、文間関係認識のための構造的アライメントを提案している。これは、異なる二つのセンテンスを比較する際に、単純な語の比較だけではなく、それぞれの文内での関係も考慮した単語アライメントを用いる手法を用いている。これにより、単純な単語アライメントよりも適正なアライメントを得ることができ、より意味的に正しい、類似・包含といった文間関係を取得することができるという結果を出している。

このように、ユーザの情報集約を助ける文間や言論間の関係抽出を行っている研究は多い。しかし、複数の文や言論にまたがる流れを考慮した事象ベースの比較はあまり行われていない。本研究では、それを軸にすることで、これまで発見が困難であったエンティティ間関係の発見を目指す。

2.1 談話構造解析

連続した文間関係の取得する手法として、談話構造解析という分野があり、先行研究がなされている。

竹井 [4] らは、日本語コーパス分析と母国語話者調査の結果から、センタリング理論とゼロ代名詞に対する考察を行っている。センタリング理論とは、Grosz [5] らによるもので、センテンスの主要エンティティの遷移を利用して、センテンス間の関係を捉えるもので、竹井らは、このセンタリング理論における重要な位置を占める主格が省略された格・ゼロ代名詞が多い日本語の文章に対して調査を行っている。

梅澤 [6] らは、センタリング理論と対象知識を用いることで、談話構造解析を行う談話構造解析システム DIA を提案している。DIA では、センタリング理論を拡張し、センテンスが参照するセンテンスを遡ってより関係の深いセンテンスを決める。

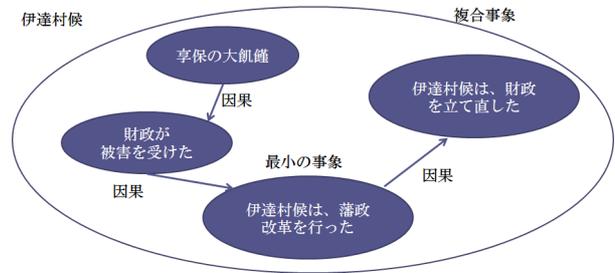


図 1 本研究における複合事象、及び最小事象の例

また、文中のエンティティの対象知識を用いることで、包含関係や上下関係を利用し、より正確なセンテンス間関係を求めるアプローチを行っていた。

また、柴田 [7] らは、隠れマルコフモデルを用いることで、センテンスのトピックを捉え、そのトピックに基づいた文間の関係・文章のトピック遷移を捉える談話構造解析の手法を提案している。

このように、談話構造解析に対するアプローチは複数あり、また使用する文書や分野によって制度の高い手法もある事が分かる。我々は、Wikipedia 情報を扱うため、この記事情報の特徴を捉えてより精度の高い手法を考えたい。

また、この談話構造解析的なアプローチを取り、得られたグラフを比較することでエンティティ間比較を行うのが本研究の目的であるため、それに適した構造や結束関係、比較手法までを考えたい。

3. 事象の定義

この節では、本稿における事象の定義を示す。

本稿では、「事象」を、事態の推移や人物の行動など、実際に起きた事実としての事象を扱う。また、本稿では、この事象を以下のように二つに分ける。

- 最小事象
- 複合事象

最小事象とは、それ以上分割できない事象の事であり、本手法で提案する事象モデルであるグラフのノードとなるものである。詳しくは後述する。

複合事象とは、複数の最小事象を組み合わせたものであり、本手法で提案する事象モデル中の部分グラフがこれに相当する。一般にユーザが事象と認識するものは、この複合事象である事が多いと考える。

これらの関係を簡単に表したものが、図 1 である。

3.1 最小事象

最小事象とは、我々が提案する事象モデルにおいて、それ以上分割できない事象である。

我々は、この最小事象になり得るのは以下のようなものであると定義した。

- 述語項
- 事象性名詞

以下ではこの二つの最小事象について述べる。

表 1 述語項の例

述語	主格	対格	キーワード
受けた	財政	被害	
行った	伊達村候	藩政改革	
立て直した	伊達村候	財政	

表 2 享保の大飢饉の内容

述語	主格	対格	キーワード
起こった	享保の大飢饉		江戸時代, 中期, 飢饉

3.1.1 述語項

述語項は、記事中に含まれる述語と、その意味を補う項によって構成されたものである。述語項構造解析に関する研究として、奈良先端科学技術大学院大学が開発した「SynCha」[10] システムなどが挙げられる。

述語に注目したのは、述語自体がある程度の意味・意図を含有していると考えためである。

本研究では簡単のため、述語の意を補う項として、主格(ガ格)と対格(ヲ格)を重視し、その他はキーワードとして扱う。

このような述語項の例として、図 1 での最小事象のうちの述語項を、表 1 に記した。

3.1.2 事象性名詞

事象性名詞とは、記事中に含まれるアンカーリンクのうち、リンク先が事件などの、事実としての記事であった場合に、その名詞自体を事象として扱うため、本研究で定義したものである。この事象性名詞をノードとして扱う場合、リンク先のリード文の 1 文を事象の内容として扱うものとする。

このような事象性名詞として、図 1 での最小事象として「享保の大改革」が例として挙げられる。この時、リンク先のリード文より、この事象の持つ意味は表 2 のようになる。

3.2 最小事象の抽出

最小事象の抽出には、形態素解析器として MeCab [11] を利用する。本稿における事象モデルは Wikipedia の記事を扱うため、MeCab の結果から、パラグラフ情報やアンカーリンクといった情報も抽出する。特にアンカーリンク抽出は、事象性名詞の抽出の他にも、固有表現の正規化に用いることができるため、重要である。

述語項の抽出は、句点、及び動詞を区切りに述語、主格、対格、キーワードを抽出する。これらは基本的に「は」「が」「を」「に」といった助詞による表層から取得するが、モデル構築において重要な主格は、該当するものが無ければ、前後の文の主格か、記事の人物を主格として補完する。

事象性名詞の抽出は、アンカーリンクの中から、カテゴリ情報で事件を持つもののみを抽出する。また、内容として、リンク先のリード文の最初の 1 文から「誰が」「何をした」事象かを抽出する。

このように抽出した最小事象は、表 1,2 のようになる。

4. 事象モデル構築

この節では、抽出した最小事象間の関係を推定し、事象モデ

表 3 TRANSITION の決定

TRANSITION	条件
CONTINUATION	$Cb(E_i, E_j) = Cb(E_j, E_k)$ であり、 $Cf(E_i)$ の最上位要素
RETAINING	$Cb(E_i, E_j) = Cb(E_j, E_k)$ であり、 $Cf(E_i)$ の最上位要素以外
SHIFTING	$Cb(E_i, E_j) \neq Cb(E_j, E_k)$
NOTHING	$Cb(E_i, E_j) = \emptyset$

ルとして関係グラフを構築するための手法を述べる。

4.1 親事象の推定

まず、各最小事象の、親となる各最小事象を推定する。本手法では原則、最小事象は親を一つだけを持つものとして定義する。親の推定には、梅澤らの拡張したセンタリング理論を基礎に行う。

センタリング理論とは、談話構造解析における手法で、Grosz [5] らが提唱したものである。文間で引き継がれたワードのうち最も上位なものを中心語として、その更新のされ方を TRANSITION として規定するものである。

梅澤 [6] らは、このセンタリング理論を拡張し、TRANSITION を求める対象文を前文だけではなく、それ以前の全ての文に広げ、そのうちで最も結束の強いものを結束のある文として選ぶ手法を提案している。この手法と、語知識を用いることで、より意味的に正しい結束関係を求めていた。

本手法では、文よりもさらに細かい単位の最小事象を扱うため、その結束関係の取得は重要である。そのため、意味的に正しい結束を求めるため、この拡張したセンタリング理論を用いる。以下に、センタリング理論の流れを記述する。なお、元手法では文 S_i を用いている部分を、本手法では最小事象 E_i に置き換えている。

- 現在の最小事象から $Cf(E_i)$ を求める
- $Cb(E_i, E_j) (i > j)$ を求める
- 事象の三つ組 $E_i, E_j, E_k (i > j > k)$ に対して、 Cb の推移 $Cb(E_j, E_k) \neq Cb(E_i, E_j)$ の値を表 1 に従って求める。
- 求めた推移のうち、最も結束の強い推移を持ち、且つ最も距離の近い E_j を親事象とする。

– 結束の強さは、以下の通りである

CONTINUATION > RETAINING >

SHIFTING > NOTHING

このように、全ての最小事象に対して親事象を求め、エッジを張るノードの組を決定する。

ここで $Cf(E_i)$ は、事象 E_i 中の中心語候補のリストで、優先順位により降順にソートされたものである。本手法では、最小事象を用いるため、この優先順位を以下のように定めた。

主格 > 対格 > それ以外のキーワード

また、 $Cb(E_i, E_j)$ は、事象 E_i 中に出現する中心語候補のうち、 $Cf(E_j) (i > j)$ に含まれる語と一致する、ソート順で最上位の要素である。なお、ルートノード事象及び全ての E_j について $Cb(E_i, E_j)$ が \emptyset だった場合、 $Cf(E_i)$ の最上位要素を $Cb(E_i)$ とする。

図 1 の例を用いると, $Cf, Cb, TRANSITION$ は以下のようになる。

- $Cf(E_1) = \{ \text{享保の大改革, 江戸時代, 中期, 飢饉} \}$
- $Cb(E_1) = \text{享保の大改革}$
- $Cf(E_2) = \{ \text{財政, 被害} \}$
- $Cb(E_2) = \text{財政 (NOTHING)}$
- 親ノード: E_1
- $Cf(E_3) = \{ \text{伊達村候, 藩政改革} \}$
- $Cb(E_3) = \text{伊達村候 (NOTHING)}$
- 親ノード: E_2
- $Cf(E_4) = \{ \text{伊達村候, 財政} \}$
- $Cb(E_4, E_2) = \text{財政 (RETAINING)}$
- $Cb(E_4, E_3) = \text{伊達村候 (CONTINUATING)}$
- 親ノード: E_3

ここで, NOTHING が多いのは, 対象を文から事象に変えたためだと考える。これは, 同一文内の事象同士は共通する Cf を持たない事が少なくないためである。

しかし, 同一文内の最小事象同士の結束は強いのが自然であるため, これを考慮する必要がある。そこで本手法では, センタリング理論では各 TRANSITION 及び距離によるスコア付けを行い, それに同一文であるかどうかのスコアを付加した上で, 最大スコアを持った事象を親事象とするなどの手法を, 今後実験を行いながら確立したい。

次に, 求めた推移のある事象同士の関係を求める手法を述べる。

4.2 親事象との関係の推定

談話構造解析において, 文間の関係として定義されているものは多数ある。

柴田 [8] らは, 二文間の結束関係として, 並列, 対比, 理由, 条件, 主題連鎖, 焦点主題連鎖, 詳細化, 理由, 原因結果, 例提示, 質問応答の 11 種類の関係を定義している。また, 梅澤 [6] らは, 詳細化, 展開, 原因結果, 逆説, 遷移, 転換, 並列, 例提示, 質問応答の 9 種類の関係を付与している。

本稿では, これらを参考に, 事象間の関係として, 以下の 8 つを与える。

- 展開, 例示, 並列, 因果, 逆説, 推移, 手段, 状況

各関係について捕捉すると, 展開とは前文で与えられた新たな話題へと話に移るもので, 話題の詳細化などもこれに含まれる。手段とは前後の事象について, 実際に行った手段や行動についての記述があれば, これに当たる。状況とは, 事象が起こった時の背景や行動した際の状況などについての記述と当該事象との関係を意味する。推移とは事象と事象とが自然に推移した関係で, 基本的に時勢順に記述される Wikipedia の人物記事においては, この関係が一般的な関係となると考える。

以上の関係を必要な関係として与えたのは, Wikipedia の人物記事の事象間関係として重要であると考えたためである。特に状況や手段は, 同一文内に記述の多いもので, また事象間比較において重要であると考えたため, 新たに追加した。

これらを, 求めた親事象との推移と, 接続詞等の表層から総合して求める。

たとえば親事象との推移が CONTINUATION であった場合, 両者の関係は, 推移である可能性が高い。また, RETAINING であった場合, 展開, 例示である事が多い。このように, 事象の推移の種類からある程度事象間の関係が得られる。また, 「理由」や「～による」, 「しかし」といった特徴的な表層表現から因果・手段・逆説等の関係を取得することもできる。

このような, 親事象との推移の種類によるスコアと, 接続詞等の表層表現によるスコアとを合計して, 関係の決定を行う。この時のスコアの配分は今後実験を行いながら適正な値を求めていきたいが, 基本的に表層表現の方が推移の種類よりもスコアに大きく影響するような配分になると考える。

5. 事象の類似

本節では, 事象が類似しているか否かを判定するための手法について考察する。

事象が類似しているか否かは, 事象の粒度によっても判定方法が異なると考えられる。

5.1 最小事象同士の類似性

最小事象同士の比較は, これ以上分割できない事象であるため, 複雑な構造を気にする必要がなく, 従来通りの語ベースによる比較で可能であると考えられる。各名詞間の関係 (主格や対格など) も考慮する必要はあるかもしれないが, 後藤 [3] らの手法等を適用することで対応できると考える。

5.2 複合事象同士の類似性

複合事象同士の比較は, 単純な語ベースの関係だけでは不十分であり, 含まれる最小事象間の関係も考慮する必要がある。つまり, 以下の条件を満たす必要がある。

- それぞれの複合事象の事象モデルが, 同じ構造を持つ
- 対応する最小事象同士が, 類似している

この構造での比較の際に, 前節まで与えた最小事象間の関係が重要な意味を持つ。

本稿で与える最小事象間の関係は, 大きく分けて次の三つに分けられる。

- 構造的な関係
展開, 例示, 並列
- 意味的な関係
因果, 逆説, 推移
- 補足的な関係
手段, 状況

構造的な関係とは, そのまま構造的な意味を持つもので, 最小事象間の関係という意味では, あまり本質的な意味のない関係であると考えられる。代わりに, これらの関係はグラフ同士の比較をする際にグラフを単純化するなどの利用ができる。例えば展開は, 特定の話題について詳細化等が行われている可能性が高く, この先を事象の流れの本質ではないと枝刈りを行う事ができると考える。また, 並列の場合は複数の並列された最小事象を縮約させることで比較が可能になる場合もあると考えられる。

意味的な関係とは, 最小事象間の意味を持ったつながりのことであり, 事象比較において本質的な関係であると考えられる。そのため, この関係については特に類似事象間で合致している必

要があると考える。

補足的な関係とは、ある最小事象と、それに対して補足的な意味を持つ最小事象との間にある関係であり、上述した二つの関係の性質を併せ持つと本稿では考える。手段や状況といった情報は補足的なものであり、本質ではないため無視する事もできるが、ユーザの意図によってはその補足的な情報が重要である事もあるため、この関係を構造の単純化に用いるのか比較に用いるのかは、柔軟に対応できなければならないと考える。

以上のような、最小事象間の関係の特徴を用いて、複合事象同士の類似しているかどうかを判断する。まず、それぞれの事象モデルのグラフを構造的な関係等により比較に適した形に変形した後、それぞれのノード間を同じ意味的關係(場合によっては補足的関係)で結ばれているかどうかを判定することで、事象間の類似があるか否かを判定できると考える。

6. 今後の課題

6.1 実装

本モデルを用いたシステムとして、以下のようなものが考えられる。

(1) 1 エンティティ・1 複合事象を入力とし、それに類似した複合事象を持つエンティティ集合を返す

(2) 2 エンティティを入力とし、両エンティティ間に類似事象が含まれるかどうかを返す

(3) 複合事象を入力とし、そのような複合事象を持つエンティティ集合を返す

(1) は、例えば Wikipedia 閲覧中に気になった箇所をピックアップしてシステムに投げることで、システムが選択文章から複合事象を抽出し、これと類似した複合事象を予め構築しておいた事象 DB から探し、ユーザに返すといった流れのシステムになる。

(3) は、応用例の一つとして考えられるもので、複合事象として「どういう理由で、どういう行動を取って、どういう結果になった」といった曖昧ながらも一定の流れに即した意味を持つエンティティの検索が行うものである。これは、概要が分かっているもののそのエンティティの名前が出てこない場合などにその答えを推定する、うるおぼえ検索として用いることができるのではないかと考える。

本研究では、目的の一つであった、記事を跨いだ情報集約のサポートとして用いることのできる (1) のシステムについての実装をまず目指す。

6.2 実験

実験として、本稿で述べたモデルの抽出手法に関して、以下のような評価実験を行う予定である。Wikipedia 人物記事から適当に数ページ選び、各ページについて以下の通りに人手で処理を行う。

(1) 選んだページから、述語項構造及び事象性名詞を人手で解析する

(2) (1) で得られた最小事象の結束関係を人手で判定する

(3) (2) で得られた結束関係にあるとされた最小事象間の関係を人手で与える

上記のようなステップで、選択した人物記事から事象モデルグラフを手動で構築する。各ステップの結果はそれぞれ、以下のものと対応している。

(1) 形態素解析等を利用した最小事象抽出の正解データ

(2) センタリング理論による結束関係推定の正解データ

(3) 結束関係及び表層表現を利用した最小事象間関係推定の正解データ

よって、これらのデータから各ステップでの抽出・推定の精度の評価を行う事ができると考えられる。また、一連の動作に関して計算時間等も観測し、Wikipedia 人物記事全体において使用に耐える手法であるかどうかの評価も行いたい。

他にも、モデルそのものの評価も必要であると考え、本モデルが発見した類似事象、及びその親エンティティの類似関係について妥当か否かをユーザに判断してもらうといった評価も行う必要があると考える。

7. まとめ

本予稿では、エンティティ間の類似関係取得のために、文章の流れを利用した、事象ベースの比較を行うという骨子のアイデアと、それを実現するための事象モデル構築に対する基本アイデアを述べた。

今後の課題として、まず実験を行いながら、提案した事象モデルの構築手法に対する妥当性を検証し、またより精度の良い構築手法を模索していきたい。

また、5 節で述べた事象間の類似性判定の具体的な手法についても考案、及び実験による検証を行う必要がある。

また、実際にシステムとして運用する際のインターフェースなど、システムとして考えた際の本手法の適用方法などについても考慮が必要であると考えられる。

文 献

- [1] 松吉俊, 村上浩司, 増田祥子, 松本裕治, 乾健太郎. 事象間関係知識の整備と類似・対立認識への応用, 情報処理学会研究報告 2008-NL-187, pp.15-22, 2008
- [2] 島田祐司. Wikipedia における同一事象検出, 京都大学大学院情報学研究科修士論文, 2010
- [3] 後藤隼人, 水野淳太, 村上浩司, 乾健太郎, 松本裕治. 文間関係認識のための構造的アライメント, NLP2010, pp.848-851, 2010
- [4] 竹井光子, 藤原美保, 相沢輝昭. センタリング理論とゼロ代名詞: 日本語コーパス分析と母語話者調査の結果から, NLP2006, 12-pp.292-295, 2006
- [5] Grosz Barbara J, Weinstein Scott, Joshi Aravind K. Centering: A Framework for Modeling the Local Coherence of Discourse, Association for Computational Linguistics, 21, pp.203-225, 1995
- [6] 梅澤俊之, 原田実. センタリング理論と対象知識に基づく談話構造解析システム DIA, 自然言語処理 18(1), pp.31-56, 2011
- [7] 柴田知秀, 黒橋禎夫. 隠れマルコフモデルによるトピックの遷移を捉えた談話構造解析, 情報処理学会論文誌 Vol.48 No.6, pp.2129-2139, 2007

- [8] 柴田知秀, 黒橋禎夫. 談話構造解析に基づくスライドの自動生成, 自然言語処理 13(3), pp.91-111, 2006
- [9] Wikipedia, <http://ja.wikipedia.org/>
- [10] SynCha: <http://cl.naist.jp/syncha/>
- [11] MeCab:Yet Another Part-of-Speechand Morphological Analyzer, <http://mecab.sourceforge.net/>