

物語文章における時系列推定の拡張 An Extension of Time Series Inference in Story Sentences

奥村 紀之† Noriyuki Okumura
瀧本 洋喜‡ Hiroki Takimoto

1. はじめに

インターネットの普及により、著作権の切れた物語文章などが公開され、文学作品に従来よりも容易に触れられる機会が増えている。青空文庫のようなサービスでは、様々な作品を一括して公開しており、ユーザにとってほどの物語が自分の感性に合っているのかを判断するのが難しくなっているため要約文生成技術が求められる[1][2]。

物語文では、論説文や新聞記事といった形式の決まった書き方をせず、ある主張に軸を置き、その前後を論拠で固めるといった構成を取らない。そのため、文末表現などを利用した要約文生成を行うことが困難である。また、物語文は登場人物(主に主人公)に関する描写や情景描写を中心として構成されており、物語の重要箇所は描画の移り変わりや他の描画との関連性によってのみ決定される。そのため、一般的な自動要約を適用することが難しい。

そこで、本研究では、物語文章に着目し、その要約文を生成する上で重要となる物語文章の時系列推定手法[3]の拡張モデルを提案している。

2. 関連研究

一般的なテキストを対象とした自動要約の手法の一つとして重要分に着目した要約手法がある。この手法では、テキスト内で重要と考えられる文を重み付けによって抽出し、要約を生成するという手法をとる。重み付けの手法としては、 $tf \cdot idf$ を用いた手法[4]や、単語同士の共起度を利用した手法[5]がある。しかし、これらの手法は文の重要度のみに着目しており、文と文の関連性については考慮されていない。そのため、代名詞に対する先行詞が存在する保証がない、文間の一貫性が悪いといった問題が発生する[6]。また、文章の内容理解を行わないため、重要分抽出によって一般的なテキスト、特に、本稿で対象とするような物語文章を要約した場合、以下の問題が考えられる。

- 複数の文で構成された場面描写を区切ってしまうことがある。
- 物語の背景となる知識が物語固有である場合が多いため、単純に重み付けをすることができない。
- 登場人物の行動記述が中心となって構成されるため、主題(物語文のタイトル)を手がかりとした重要分の判定が出来ない。
- 意図的に錯時[7]された文章を要約した場合、時系列が前後した要約が作成されてしまう。

自動要約に関する研究は、論説文や新聞記事を対象としてきたものが多い。そこで、上記の4つの問題を物語固有の特徴としてとらえ、要約生成を行う研究も進められている。物語文における場所や人物、時間に着目し、物語のシーン分割を行う手法[8]では、場所、登場人物、時間を手がかりとして、物語の場面の変わり目を抽出している。また、登場人物の情報を体系化する手法[9]は、同じ場面に登場する人物同士には何らかの関係があると仮定し、登場人物に着目したモデル化を行っている。これらの研究は、物語文章に見られる独特の性質を利用することで、より効果的に物語文章の解析を進めている。

別の観点からのアプローチとして、物語の構造に着目した自動要約の研究も行われている。話題間の接合に留意し、整合がとれた要約の生成する手法[10]では、物語を登場する話題ごとにいくつか分割し、分割された話題と話題の間を補完する文章を抽出することで、要約文生成を行っている。感情に着目した手法[11]では、登場人物の感情の変化が示された部分は他の部分に比べて重要であるという仮定に基づき、物語の登場人物の感情表現を抽出することで要約を生成している。人のテキスト理解モデルによる手法[12]では、人間がどのようにテキストを理解、記憶し、どのような箇所に注目して要約を行うかという点に着目し、談話理解モデルに基づいた物語の重要分抽出を行っている。

このように、物語独特の特徴に着目した自動要約の研究が行われている。一方本研究では、要約文生成の際に、文章が持つ時間に関する情報を適宜推定し、物語文の時系列推定を行うことが関連研究とは一線を画している。

なお、本研究における基幹システムとして、我々は時間に関する知識ベースを用いた時間判断システムを提案している[13]。時間判断システムでは、単語、あるいは文章がどの時間帯、どの時期に所属するものかを判定することが可能であるため、本研究における物語文章の時系列推定の根幹を形成している。

3. 物語文章からの時系列推定による自動要約

物語文章の著者は、本来時系列順に記載される出来事を、意図的に操作する場合がある。本研究では、時系列の操作が行われたことを明らかにし、時系列が順当に進行したことを明確に判断できることを合わせて時系列推定と定義する。すなわち、出来事のみ鋭角な時間や日時を抽出することについては本研究の対象外となる。

物語の本来の時間進行とは別に発生する時系列の操作は、著者の意図する効果(狙い)によって様々である。数日程度の期間を記述しないことによる時間の跳躍から、以前の出来事を振り返ったり、将来について想像しているようなことまで多岐にわたる。しかし、著者が読者に何らかの印象を与えるための操作として時系列の入れ替えが行われる以上、時系列の操作は文章中の語句を通じて読者に理解でき

† 香川高等専門学校情報工学科
Kagawa National College of Technology, Department of
Informattion Engineering

‡ 東京工業大学精密工学研究所
Tokyo Institute of Technology, Precision and Intelligent
Laboratory

る表現で示されていると予測できる。ここで、著者が時系列操作を読者に知らせる表現と、物語本来の時間進行を合わせて時系列情報と定義する。また、時系列情報を表現する語句は、時系列推定を行う上で手がかりとなる語句で時語と呼ぶ。

3.1 システムの全体構成

図1に時系列推定を組み込んだ文章要約システムの構成を示す。

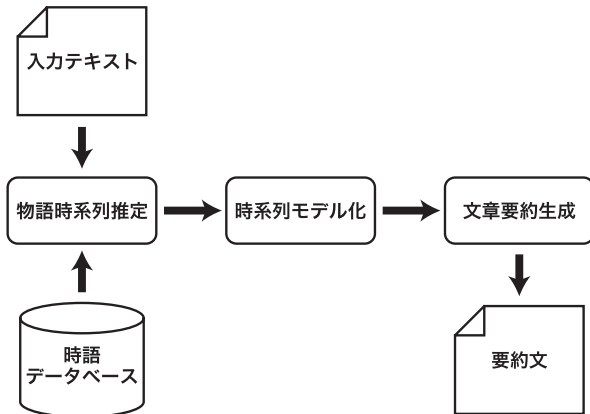


図1. 時系列推定を組み込んだ文章要約システム

本システムでは、まず第一に、一つの文章が持つ時系列情報は一つの時系列情報に帰着する、あるいは時系列情報を持たないのいずれかであると仮定する。物語文章における各文に対して、時系列情報に帰着できるか、帰着できるのであればどのような時系列情報であるのかを推定する。推定の際には、時節に述べる時語データベースを参照し、文に含まれる時語を手がかりとして時系列推定を進める。この手順を時系列推定と呼ぶ。

第二に、物語全体の時系列の流れのモデル化を行う。物語の中には、推定した結果時系列情報を持たない文章も数多く存在すると考えられる。そこで、時系列情報を持たない文の時系列情報を、直前、あるいは直後の時系列情報を持った文の時系列情報から推定し、物語全体の時系列の流れをモデル化する。

最後に、実際の要約部分は本稿の対象外ではあるが、時系列推定、および時系列のモデル化によって得られたモデルを解析し、連続した時系列の区切れ目や、時系列操作が行われた区間などを手がかりに、時系列を考慮した要約文の生成につなげる。

3.2 時語データベース

本研究では、時間に関する情報を保持した時語データベースとして、時間帯単位、年月日単位、季節単位、人生単位の四種類に分類した代表的な時語 511 語を使用する。時語データベースは、103 名の被験者に、「時を表現する語句」と「時を連想できる語句」、「人に一生の中のある期間を表現する語句」の三項目に対してアンケート調査を実施しサンプルを収集した。収集したサンプルを 5 名の被験者に判定させ、4 名以上が時間に関係した語句であると評価したものを時語データベースとして採用した。以下に、4 種の単位の定義と例を示す。

- ・時間帯単位時語

一日の中のある時間帯を表す時語であり、二種計 110 語を登録している。

(例: 「基準時語」:朝, 夜 「連続時語」:始発, 下校)

- ・年月日単位時語

一年や一日単位の移り変わりを表す時語であり、三種計 60 語を登録している。

(例: 「基準時語」:日, 年 「相対時語」:今日, 再来月 「週時語」:火曜日, 木曜日)

- ・季節単位時語

一年の季節のどこかを表す時語であり、三種計 257 語を登録している。

(例: 「基準時語」:春, 夏 「指定時語」:春分, 七夕 「連想時語」:海, コスモス)

- ・人生単位時語

人の一生のある期間を表す語句であり、84 語を登録している。

(例: 高校生, 生年)

4. 物語時系列推定とモデル化

本節では、時語を手がかりとした自動要約の基盤をなすシステムである時系列推定システムの構成を図2に示す。時系列推定システムは、物語文章を入力として与えると、各文の時系列情報を推定し出力する。時系列情報は、一文に対して時間帯、年月日、季節、人生の四つの異なる情報を推定する。時系列情報を推定する手がかりとして時語データベースを参照し、適宜概念ベースによる未知語処理を行う。

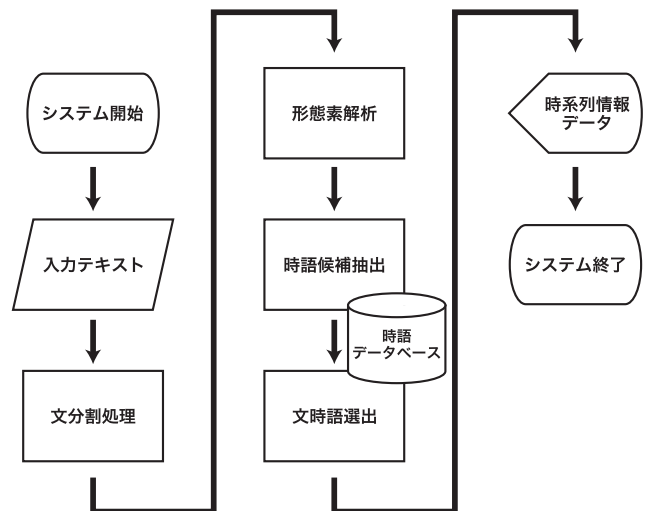


図2. 時系列推定システムの流れ

本システムを構成するための先行研究として構築している、概念ベース、関連度計算方式、未知語処理について記述したのち、システムの詳細について述べる。

4.1 概念ベース

時語データベースに登録されている語数は 511 語と少なく、時語データベースに存在しない表現も物語文章には数多く存在している。そのため、時語データベースを柔軟に拡張し参照する必要がある。本研究では、概念ベースと関連度計算方式による未知語処理によって、時語データベースの拡張を実現している。本節では未知語処理の基盤となる概念ベースについて述べる。

概念ベース[14]は、電子化辞書や電子化新聞から機械的に構築された大規模知識ベースであり、およそ20万語の表記集合で構成されている。概念ベースにおいて、任意の概念 A は、概念の意味や特徴を表す属性 a_i と、 a_i が概念 A に対してどの程度重要であるかを示す重み w_i の対の集合で定義されている。

4.2 関連度計算方式

関連度[15]とは、任意の二つの概念 A, B の関連の強さを概念ベースによって定量化した値である。関連度計算方式には意味関連度、共起関連度、漢字関連度など種々の手法が提案されているが、本研究では重み比率付き意味関連度計算方式を用いる。

任意の概念 A, B について、それぞれ一次属性を a_i, b_j とし、対応する重みを u_i, v_j とする。また、概念 A, B の属性数を L 個、 M 個 ($L < M$) とする。

$$A = \{(a_i, u_i) | i = 1 \sim L\}$$

$$B = \{(b_j, v_j) | j = 1 \sim M\}$$

このとき、概念 A, B の一致度 $MatchWR(A, B)$ を以下の式で定義する。

$$MatchWR(A, B) = \sum_{a_i=b_j} \min(u_i, v_j)$$

$$\min(\alpha, \beta) = \begin{cases} \alpha (\beta \geq \alpha) \\ \beta (\alpha > \beta) \end{cases}$$

$$\max(\alpha, \beta) = \begin{cases} \alpha (\beta \leq \alpha) \\ \beta (\alpha < \beta) \end{cases}$$

概念 A, B の属性 a_i, b_j に対し、 $a_i = b_j$ (概念 A, B に共通する属性がある) となる属性があった場合、共通する属性の重みの共通部分、つまり、重みの小さい分だけ有効に一致すると考え、このように一致度を定義する。定義から明らかな様に両概念の属性と重みが完全に一致する場合に一致度は1.0となる。

概念 A と概念 B の関連度 $MR(A, B)$ は、

1. 概念 A と概念 B の属性を重み順に上位 t 個抽出する。
2. 属性の少ない方の概念を A とし、概念 A の属性を基準とする。

$$A = \{(a_1, w_1), (a_2, w_2), \dots, (a_L, w_L)\}$$

3. 属性どうしが完全に一致する属性 ($a_i = b_j$) は別扱いとする。すなわち、 $\max(u_i, v_j)$ となる a_i または b_j の重みを $\max(u_i, v_j) - \min(u_i, v_j)$ とし、その重みを u'_i あるいは v'_j とする。
4. 完全に一致する属性を除いた概念 A' と概念 B' は、

$$A' = \{(a'_1, u'_1), \dots, (a'_t, u'_t), \dots, (a'_L, u'_L)\}$$

$$B' = \{(b'_1, v'_1), \dots, (b'_t, v'_t), \dots, (b'_M, v'_M)\}$$

のように表せる。ここで、概念 A' と概念 B' には完全に一致する属性が存在しないので、類似する属性について一致度を計算する。一次属性の一致度により全体の一致度の和が最大となるように対応を決め、定まった各属性の一致度とそれぞれの属性の重みを用い、関連度を求める。

5. 完全一致する属性が α 個あったとすると、概念 A と概念 B の意味関連度を以下に示す式で定義する。

$$MR(A, B) = \sum_{i=1}^{\alpha} \{\max(u_i, v_j) - \min(u_i, v_j)\} + \sum_{i=1}^{L+\alpha} MatchWR(a'_i, b'_j) \times \frac{u'_i + v'_j}{2} \times \frac{\min(u'_i, v'_j)}{\max(u'_i, v'_j)}$$

このように関連度を定義しているため、関連度の妥当性は如何に正しい属性が付与され、適切な重みが与えられているかに大きく依存する。すなわち、概念 A と概念 B に対し重要性が高く、かつ、その重みが近似するような場合、関連度が大きくなる。

4.3 未知語処理

未知語処理とは、時間判断システムなど常識的判断システムにおいて、個別に定義した知識ベースを概念ベースによって自動拡張するための手法である。未知語処理を導入することにより、各種専用知識ベースの規模を最小限に抑え、語彙の拡大は概念ベースによって対応する。未知語処理では、時語知識ベース内の511語と時系列推定の対象となっている語を関連度計算法によって比較し、時語知識ベースに存在しない語がどの語に対応づけられるべきかを推定している。

未知語処理手法は、未知語 X との関連度が最も大きい既知語に未知語 X を帰着させ、その既知語が表現する時系列情報を、未知語 X の表現する時系列情報とする。ただし、概念または関連度の性質上、ある語とある語の間に関係がないと考えられる場合であっても、小さな値の関連度が出力されてしまう。そのため、単に「最も大きな関連度を持つ既知語」という判断基準により処理を行った場合、全ての既知語に関係がなかったとしても、未知語 X と関連の弱い既知語に帰着する恐れがある。そこで、判断基準にしきい値 Thr を設け、最も大きな関連度がしきい値 Thr 以下であった場合には、未知語 X をその既知語に帰着させず、時系列情報を持たない語であると推定する。

4.4 文時語推定処理

本研究における前提条件として、一つの文が帰着すべき時系列情報は一つであるという条件がある。しかし、文は多くの単語や表現から構成されており、一文に含まれる複数の語が時系列情報を持つ可能性はある。また、複数の時系列情報を持つ語が取得された場合、どの語が優先的に使用されるべきかは文の内容によって様々であるため、一位に優先度を設定することが出来ない。そこで、推定する文章と最も関連の高い時系列情報を、文全体の時語として導出する手法をとる。

1. 入力文中に出現する単語を、未知語処理を用いて既知の時語に帰着し、時語候補群を得る。
2. 得られた時語候補群と入力文中の単語との関連度を計算し、その和を求めて時語候補群と入力文との関連度計算を行う。
3. 入力文と最も関連の高かった時語候補から得られる時系列情報を入力文が表す時語候補とする。ただし、未知語処理の場合同様、関連度計算の性質上、ある語とある語の間に関係がないと考えられる場合でも、小さな値の関連度が出力されてしまう。その為、判断基準にしきい値 Ths を設け、入力文と最も関連の高い時語との関連度がこのしきい値 Ths 以下であった場合は、

入力文と時語を関連付けず、時系列情報を持たない文であると推定する。

なお、文事語推定処理に使用するパラメータを設定するための予備実験を行っている。予備実験では、物語文章からランダムに取得した347文に対して未知語処理と文事語推定処理を行った結果について3名の被験者に判定させ、3名全てが関連していると評価した関連度の下限值と、2名が関連していると評価した関連度の上限を求めその中間値を採用している。その結果から、Thrを0.7、Thsを0.75と設定した。

4.5 時系列のモデル化

物語中のある点で出現した時系列情報は、時系列の操作がない限り、それ以降の文にも時間的に似通った時系列情報が出現すると考えられる。そこで、物語中の時系列情報を利用した時系列の補完を行うシステムとして物語時系列のモデル化を構築している。このシステムは、文事系列推定済みの物語文章を入力とし、注目している文の近傍の時系列情報を用いて、時系列を持たない文の時系列を推定し時系列の明らかな文から推定された時系列の信憑性を向上させる。時系列情報は、時語データベースの4種の単位それぞれを別の情報として扱う。

以上の仕様を満たす長さLの文から構成された物語におけるある文xの時系列補完 Tl_x を以下の式で行う

$$Tl_x = \frac{\sum_{i=x}^{x+f} t_i}{f+1} \quad f = L \times p$$

ここで t_i は時系列推定によって推定された文iの時系列情報、pは物語全体の文鳥の何パーセントを補完のための文として利用するかを決定する幅である。ある文xの時系列を補完する際に、x以降x+fまでの文章の時系列情報を総和し、その平均値をxに対する時系列補完の結果として扱う。

5. 評価実験

本節では、4.5節に示した時系列のモデル化システムの評価実験として青空文庫[16]から入手できる作品の中で、「麦わら帽子」を対象として評価を行う。麦わら帽子に対しアンケートによって理想的な季節単位時系列モデルを作成し、理想的なモデルと推定結果との比較検討を行う。4.5節の式におけるpを3, 4, 5とし、pに対する補完結果を評価した。

麦わら帽子に対する時系列推定とモデル化の結果を図3に示す。図3の上のグラフには麦わら帽子の各文の時系列推定の結果をグラフにしている。縦軸は季節の時間経過を表し、最上部が元旦、最下部が大晦日に当たる。横軸は物語先頭からの分数を示す。縦軸の1ブロックはおよそ1ヶ月であり、横軸の1ブロックは20文を表している。グラフにおける棒は、文が持っていた時系列の範囲を可視化したものである。

図3の下グラフは上のグラフに示した時系列推定結果を4.5節に示した時系列補完を用いてモデル化したものである。縦軸や横軸は上のグラフと同様に、それぞれ季節の

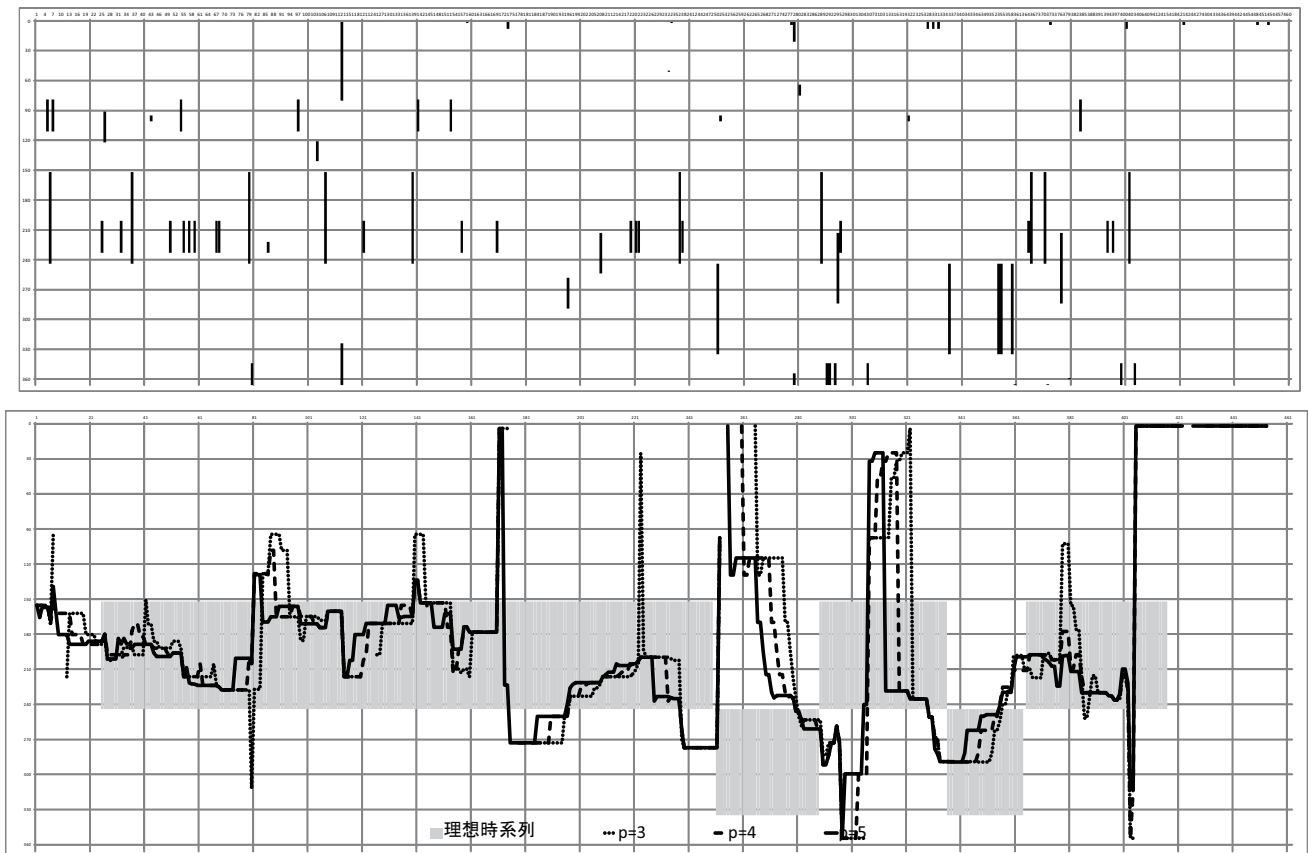


図3. 麦わら帽子の時系列モデル

時間経過と物語先頭からの文数である。帯状で示された部分は人の感性に近い季節時系列モデルである。折れ線は、 p を 3, 4, 5 としたときの時系列モデルである。

表 1 には、麦わら帽子の季節時系列において、 p を 3, 4, 5 としたときのモデル化の制度を示している。麦藁帽子は全部で 461 文であり、その内所属する時系列が明らかな文が 390 文、所属する時系列が不明瞭な文が 71 文あった。表には、モデル化の結果時系列情報が付与された文数、時系列が明らかな文の内正しい範囲の時系列を推定できていた文数、時系列が不明瞭な文の内時系列情報を付与しなかった文数を示している。

表 1. 麦わら帽子の季節時系列モデルにおける精度

文の種類 \ p	3	4	5
時系列情報が補完された文数	403	435	448
時系列情報が明らかな文(390文)での正答文数	227	252	267
時系列情報が不明瞭な文(71文)での正答文数	25	16	11
一文の補完に使われる文の平均	2.3	3.1	3.9
正答文数の合計	252	268	278
精度	0.547	0.581	0.603

6. 考察

作成したシステムの評価結果より、本研究に対する考察を行う。

6.1 時系列モデル化の有効性と課題

4 節に示した時系列モデル化の結果の一例は図 3 に示す通りである。推定前は物語全体に疎らに存在していた時系列情報を利用し補完を行うことで、時系列情報を直接持っていなかった文に対しても時系列情報を付与し、物語全体の時系列をモデル化できていることが分かる。また、時系列推定の時点で誤った情報を推定してしまった文に対しても、時系列情報の補完を行うことで誤りを訂正できることが確認できる。

また、80 文目付近にピークとして発生しているノイズが、フレーム幅 p を変更することで減衰し、 $p = 5$ のフレーム幅を使ったモデル化では、ノイズがほぼ消えていることが確認できる。正しい推定情報の中に一部誤った推定情報が入った場合は、フレーム幅の調整で誤情報の訂正ができると考えられる。

一方で、290 文目付近や 400 文目の様に、誤った推定情報が連続している場合は、時系列モデル化による補完でも誤りを補正できないことも確認できる。モデル化の精度を高めるためには、モデル化時のフレーム幅の最適化だけでなく、モデル化の前処理である時系列推定の精度向上も必要である。

6.2 p と生成モデルの関係

フレーム幅を $p = 3, 4, 5$ と変更したとき、生成されるモデルを数値として纏めたものが表 1 である。表中の情報が補完された文数の推移をみると、 p を大きくすれば大きくするほど補完される文数が多くなる。これは、 f を大きくすると一つの時系列情報がより広範囲の補完で利用されるためである。

フレーム幅が大きくなると、一つの文の時系列を補完する際に使われる時系列情報の数が増えていることから、より平滑化された時系列情報が補完されるようになると考えられる。平滑化された時系列情報は 80 文目付近に現れている様な突発的なノイズを無視することが出来、結果として時系列が明らかである文において、より正しい補完が行われていることが分かる。

一方、時系列情報が不明瞭な文に対しては時系列情報が補完されないことが望ましい。 f が大きくなれば一つの時系列情報が影響する範囲が広がるため、時系列情報が不明瞭な文に対しても情報が補完されてしまうと考えられる。表 1 の結果からも、フレーム幅が大きくなるごとに、時系列情報が不明瞭な文での正答文数が少なくなっていることが読み取れる。

時系列情報が明らかな文と不明瞭な文、それぞれに対する正答文数を合計すると、麦藁帽子では $p = 5$ の場合が最も良い精度を示した。これは、麦藁帽子の全 461 文の内、読み進めれば時系列が明らかである文が 390 文で、時系列が不明瞭な文の 71 文より圧倒的に多い事が原因だと考えられる。元々時系列が不明瞭な文が少なかった為、 p が大きくなることによる悪影響が少なかったと考えられる。

7. おわりに

本稿では、著者による意図的な錯時操作に着目し、時間的因果性を損なわないような要約手法を提案した。また、その手法の一部として物語時系列推定手法と物語時系列モデル化手法を提案し、システムの構築と評価を行った。今回構築したシステムにより、物語文章から時系列を推定し、著者による錯時や物語本来の時間進行が発生した部分を特定することができると考えた。

今回提案した手法は大きく分けて、時系列の推定、モデル化、要約作成の三段階に分割することができる。一段階目である時系列の推定では、推定する時系列を時間帯、年月日、季節、人生の四つの単位に分け、それぞれに対して時語データベースを作成し、文が表す時系列の推定を行った。二段階目である時系列のモデル化では、文自体から時系列情報を抽出できない場合でも、文の近傍に存在する時系列情報を使って時系列情報を補完した。

今後は、評価対象の物語文章を増やし、物語のジャンルによる違いや、各時系列情報の要約文に対する有用性を検証していく。

謝辞

本研究の一部は科研費 (23720222) の助成を受けたものである。

参考文献

- [1] 砂山渡, 谷内田正彦, 文章要約のための特徴キーワード発見による重要文抽出, 情報処理学会研究報告 NL-135, pp.103-110, 2000
- [2] 近藤恵子, 奥村学, 言い替えを使用した要約の手法, 自然言語処理研究会報, Vol.96, No.114, 137-142, 1996
- [3] 瀧本洋喜, 奥村紀之, 物語文章における時系列情報の抽出, 信学技報, Vol.111, No.119, NLC2011-1, pp.1-6, 2011.
- [4] 平尾努, 前田英作, 松本裕治, Support Vector Machineによる重要分抽出情報処理学会研究報告, 63-16, 2001.
- [5] 砂山渡, 谷内田正彦, 文章要約のための特徴キーワード発見による重要文抽出法—展望台システム—, 情報処理学会研究報告, NL-135, 2000.
- [6] 奥村学, 難波英嗣, テキスト自動要約, オーム社, 2005
- [7] ジェラルド・ジュネット, 物語のディスクール—方法論の試み, 水声社, 1985
- [8] 小林聡, 場・時・人に着目した物語のシーン分割手法, 音声言語情報処理, No.47, pp.25-30, 2007
- [9] 馬場こづえ, 藤井敦, 小説テキストを対象とした人物情報の抽出と体系化, 言語処理学会年次大会発表論文集, No.13, pp.574-577, 2007
- [10] 横野光, 整合性を考慮した物語要約システムの構築, 自然言語処理, Vol.15, No.5, pp.45-71, 2008
- [11] 横野光, 登場人物の感情表現に着目した物語要約, 言語処理学会年次大会発表論文集, Vol.13, pp.1141-1144, 2007
- [12] 熊坂守晃, 村上聡, 榎津秀次, 古宮誠一, 談話理解モデルに基づく物語における重要文の自動抽出法, 電子情報通信学会技術研究報, Vol.107, No.540, pp.43-48, 2008
- [13] 土屋誠司, 渡部広一, 河岡司, 連想メカニズムを用いた時間判断手法の有効性の検証, 自然言語処理研究会報告, No.73, pp.113-118, 2005
- [14] 奥村紀之, 土屋誠司, 渡部広一, 河岡司, 概念間の関連度計算のための大規模概念ベースの構築, 自然言語処理, Vol.14, No.5, p.41-64, 2007
- [15] 渡部広一, 奥村紀之, 河岡司, 概念の意味属性と共起情報を用いた関連度計算方式, 同志社大学理工学研究報告第48巻第3号
- [16] 青空文庫 <http://www.aozora.gr.jp/index.html>