

場・時・人に着目した物語のシーン分割手法

小林 聡

島根大学 総合情報処理センター

skoba@ipc.shimane-u.ac.jp

プロップに代表される、構造主義における物語の構造分析の研究では、登場人物に重点を置いて分析がなされていた。しかし、物語の構造分析の基礎といえるシーン分割についての計算機を利用した研究は、従来の場所と時間によるシーンの定義に沿って行なわれており、両者には齟齬がある。構造主義における物語の構造分析によって得られた知見を有効に利用するためには、両者の齟齬を解消する必要がある。本研究では、その端緒として、場所、時間、登場人物に着目し、物語のシーン分割を試みた。作成したシステムにより、物語に現われる場所、時間、登場人物を示す語句を推定し、それらの文間の異り数に基づいてペナルティを求め、ある文境界がシーン境界の候補となるか否かを判断した。その結果、F 尺度の最高値で、調査用データでは 54、評価用データでは 56 という結果を得た。

Scene Segmentation Method for Folktales based on Place, Time and Cast

Satoshi KOBAYASHI

General Information Processing Center, Shimane University

Structural analysis of stories in structuralism, represented by Plopp, focused on characters. However, most scene segmentation methods performed by computer are based on traditional place and time scene definitions. There are mismatches between them. To use the results of analysis in structuralism, we need a bridge between them. As the first step, the author attempted scene segmentation for folktales based on description of place, time and cast. The scene segmentation system estimated words that designated place, time or cast. Then the system calculated penalty from differences of words for place, time and cast between sentences. After that, based on the penalty, the system judged a sentence boundary appropriate as a scene boundary. As a result, the system achieved 54 for the training set and 56 for the test set in F-measure.

1 はじめに

人間がどのような方法で、またどのような形式で世界を理解しているのかを知ることは、人工知能やその他の研究領域において重要な課題である。対して、民俗学もしくは神話学においては、神話などを含む説話には、人々の世界観が現れているという考え方がある。ならば、人間が説話を、方法、形式両面からどのように理解するのかを研究することは、すなわち人間がどのように世界を理解しているのかを知る端緒となる。そのためには、計算機を用いたモデルの検討が有用であろう。

ここで、世界を理解する際に重要と思われる能力の一つを、バルトにならい仮に「分節」化する能力と呼ぶ [1]。これは、体験や記録を意味的あるいは内容的な視点を持って、何が一まとまりのモノゴトであり、何が別個のモノゴトであるかを認識・識別する能力であるとす。この、「分節」化する能力とは、単に体験や記録を分割する能力ではない。分割は必要ではあるが、分割された単位が上位の単位から見てどのような位置や役割あるいは意味を持っているのかを認識・識別し、また分割された単位間の関係を理解することにより、体験や記録は「分節」化がなされ、また個々の単位は「分節」となると考える。この際、記録からの重要箇所の抽出や要約などの技術の応用も可能であろう [2]。

また、物語の構造分析に関する研究には、プロップに始まり、グレマスやバルトらによる、構造主義における一連の研究がある。

プロップはロシアの魔法昔話を基に、一般に「プロップのファンクション」と呼ばれる、物語において継起する出来事の列を登場人物の機能 (ファンクション) に着目して整理したものを示している [3]。プロップは「プロップのファンクション」を求めるあたって、以下のような仮定を置いている。

1. 昔話の恒常的な不変の要素となっているのは、登場人物たちの機能である。その際、これらの機能が、どの人物によって、またどのような仕方、実現されるかは、関与性を持たない。これらの機能が、昔話の根本的な構成部分である。
2. 魔法昔話に認められる機能の数は、限られている。

表 1: バルトの機能体

機能体:
機能体:
枢軸機能体: 物語 (または物語の断片) の真の蝶番となるもの。それが指示する行為が、物語内容の続きに対して、因果的な二者択一を開始 (または維持、または閉止) するもの。
触媒: 蝶番をへだてている物語空間を《埋める》だけのもの。
指標:
指標: ある性格、ある感情、ある雰囲気 (たとえば疑惑のそれ)、ある哲学に関する固有の意味での指標。
情報提供者: 同定し、時間的空間的に位置づける役目を果たす情報。

なお、ここでいう「機能」とは、プロップは以下のようなものと定義している。

[そこで、ここでは] 機能という用語を、登場人物の行為で、しかも、筋=出来事全体の展開過程にとって当の行為がもちうる意義〔位置〕という観点から規程された登場人物の行為、というふうに解することにします。

また、グレマスは、物語は登場人物間の相互作用によって表わすことができるとし、6種類の登場人物を用いた行為項モデルとして、それを示している。[4]。

バルトは「物語内容のある種の線分を単位に変えるのは、その線分の機能的性格である」と述べ、表1のような一連の機能体等を提案している [1]。バルトの枢軸機能体は、言わば行為あるいは出来事であるが、そのような行為を行なう者という捉え方をすれば、ほぼ登場人物に言及しているとみなせる。

このように、プロップに端を発する、構造主義における物語の構造分析の研究は、多くが登場人物ないし登場人物の行為に着目し、物語の単位等を提唱している。

対して、演劇や映画、TVドラマ等の考え方に則した、物語の単位の捉え方がある。これは一般にシーンあるいは場面と呼ばれる。Stamらは、シーンを以下のように定義している [5]。

シーン (SCENE) — 裂け目や中断がないと感ぜられる空間・時間的な継続性。

プロップあるいは他の構造主義における物語研究者によるものと異なり、登場人物はここには表われていない。つまり、プロップらによる物語の構造分析の研究と、従来のシーンの概念あるいはシーン分割手法とでは、登場人物に着目しているか否かの点において齟齬がある。

しかし、例えば「プロップのファンクション」と物語とを対応させる等、構造主義における物語分析によって得られた知見を工学的に利用することを考えると、物語のシーン等も登場人物を考慮し、例えば「プロップのファンクション」に対応した形で分割がなされるのが望ましい。

そこで本研究では、計算機による物語理解のために物語を「分節」化すること、およびプロップに端を発する構造主義における物語分析によって得られた知見の工学的応用の端緒として、場所、時間、登場人物に基づいて物語のシーン境界を推定する予備実験を行なった。

2 実験

2.1 準備

本研究では、物語として日本の昔話を対象とした [6]。掲載された 101 話の内、80 話を用いて、表 2 に示すシーン境界の基準、および表 3 に示す登場人物の定義を求めた。表 2 の「使用」欄に○印が付いているものが、本研究において考慮した基準である。ただし、今回の実験では表 3 の条件は使用していない。

また、これらの基準の妥当性を確かめるため、2 名の被験者により、これらの基準に基づいてシーン境界ラベルを人手により与え、ラベルの一致率を求めた。その結果を表 4 に示す。物語 b で一致率が低いが、これは登場人物の入れ替わりが頻繁であったため、どのタイミングで区切りとするかの判断にズレが生じたものと思われる。しかし、シーン境界の基準としては概ね妥当であると判断した。

シーン境界の推定に関する実験には、この 80 話内の 11 話を調査用として、この 80 話以外の 4 話を評価用として用いた。

シーン境界の推定に先立ち、物語テキストの形態素解析および構文解析を行なった。形態素解析には JUMAN ver. 5.1[7]、構文解析には KNP ver. 2.0 を使用した [8]。また、名詞および一部の動詞については、場所等

表 2: シーン境界の基準

条件	使用
a: 舞台によるもの	
a-1: 場所 (舞台) が変更される箇所	○
a-2: 舞台の状態が変化する箇所	×
b: 時間によるもの	
b-1: 物語中の時間が不連続となる箇所	○
c: 登場人物によるもの	
c-1: 舞台上にいる人物に増減が生じる箇所	○
c-2: 物語の中心として語られている人物が変わる箇所	○
d: 語り手の視点によるもの	
d-1: 読み手に対する状況説明の開始点・終了点	△
d-2: 物語のまとめ (含むエピローグ) や、訓示の開始点・終了点	△

表 3: 登場人物の定義

a: セリフがある者
b: 他の登場人物や舞台を対象にした行動を起こしている者

の概念として有効と思われる概念情報を、NTT 日本語彙体系 CD-ROM 版を用いて付加している [9]。NTT 日本語語彙体系 CD-ROM 版の検索には、EB ライブラリ [10] と RubyEB[11] を用いている。

2.2 場・時・人候補語句の抽出

次に、場所、時間、登場人物の候補となる語句の抽出を行なった。

まず場所を示す語句については、助詞のついた名詞から、準備の段階で場所に関する概念を含む語句であるとされたものを場所候補とした。

表 4: シーン境界ラベル一致率

	区切箇所		一致率 [%]
	一致	不一致	
物語 a	24	0	100
物語 b	18	6	75
物語 c	30	2	94
物語 d	42	2	95
平均	—	—	91

時間を示す語句については、名詞のうち、準備の段階で時間に関する概念を含む語句であるとされたもの、および接続詞等の手がかり語を時間候補とした。

なお、JUMAN/KNPにおいて相対名詞とされる語については、場所候補にも時間候補にもなりうる。今回は、相対名詞に先行する語句およびその品詞や活用情報に基づき、それらの相対名詞を場所候補とするか、時間候補とするかを決定した。

登場人物を示す語句については、助詞「が」あるいは「は」が続く語句、および人に付く接尾辞（「さま」等）が続く語句を登場人物候補とした。ただし、名詞の概念として抽象的なものを示す語句や時間を示す語句、また形式名詞等は登場人物候補から除いた。

なお、場所候補、時間候補、登場人物候補のいずれも、一度抽出を行なったのちに再度物語テキストを走査し、先に抽出から漏れた各々の候補を抽出している。

この結果の、場所、時間、人物に関する語句の、のべ出現数に対する抽出性能を表5に示す。F尺度で、場所候補については72程度、時間候補と登場人物候補については86程度の性能となっている。

2.3 シーン境界推定実験

以上のような各候補の抽出結果に基づき、物語のシーン境界を推定する実験を行なった。

基本となる推定方法は以下のようなものである。

まず、シーンの開始文から、場所候補、時間候補、登場人物候補の語句を、各々のプールに蓄積する。各々のプール内の候補は、各々の候補が空でない二文分保持される。また、処理対象となる新しい入力文毎に、その文に含まれる場所候補、時間候補、登場人物候補の語句と、各々のプールに含まれる語句との異なり数に基づいてペナルティを計算する。特に、時間については、接続詞等の手がかり語が出現した際にもペナルティを与えている。また登場人物については、既知/未知によってもペナルティを変えている。これらの、ペナルティを与える条件やペナルティの値は経験的に決定した。結果として得られるペナルティが閾値を越えていれば、入力文が新しいシーンの冒頭であると推定する。

また、場所候補、時間候補、登場人物候補についてのペナルティを2つ、あるいは3つ組み合わせてシーン境界を推定する場合には、各々のペナルティの重み

付きの和を最終的なペナルティとした。例えば、3つを組み合わせる場合は以下ようになる。なお、この際、場所候補、時間候補、登場人物候補に関するペナルティは正規化している。

$$\begin{aligned} \text{ペナルティ}_{\text{最終}} &= 0.13 * \text{ペナルティ}_{\text{場所}} + \\ &0.55 * \text{ペナルティ}_{\text{時間}} + \\ &0.32 * \text{ペナルティ}_{\text{登場人物}} \end{aligned}$$

このような方法によって得られたシーン境界の結果を表6に示す。

なお表中の「改良」とは、上記方法によりシーン境界と推定された箇所に対して、場所候補、時間候補、登場人物候補の少なくとも1つが空でない2文前、もしくは直前のシーン区切の近い方から各プールの再構築し、ペナルティを再計算してシーン境界であるかを推定しなおした結果である。

また「時&人(2)」は、処理対象となった各文において、直前のシーン境界から各々のプールの構築し、時間と登場人物のみに対してペナルティを計算し、そのペナルティに基づいてその文がシーン境界であるかを推定した結果である。「改良(2)」は、処理対象となった各文において、直前のシーン境界から各々のプールの構築し、時間と登場人物のみに対してペナルティを計算し、一旦シーン境界の推定を行なう。その後、シーン境界候補の前後各1文の登場人物候補を比較し、両者に同じ候補が特定の助詞を伴って現れている場合には、そこをシーン境界候補から除いた結果である。

結果として、シーン境界推定の性能は、場所候補、時間候補、登場人物候補を個別に利用した場合、F尺度でおおむね45程度となった。また、場所候補、時間候補、登場人物候補のうちの2つを組み合わせた場合、F尺度でおおむね47程度となった。3つを組み合わせた場合、F尺度で53程度となった。

3 考察

対象が異なるため今回の結果とは単純には比較できないが、場所候補、時間候補、登場人物候補の3つを組み合わせた結果は、F尺度で、福田ら[12]の結果より高いが、柴田ら[13]の結果より低い値となっている。

表 5: 語句抽出性能

対象	調査用						評価用					
	正解	脱落	挿入	Recall	Precision	F 尺度	正解	脱落	挿入	Recall	Precision	F 尺度
場	263	50	141	84.0	65.1	73.4	84	6	61	93.3	57.9	71.5
時	262	7	70	97.4	78.9	87.2	75	4	21	94.9	78.1	85.7
人	629	46	146	93.2	81.2	86.8	161	7	45	95.8	78.2	86.1

表 6: シーン境界推定性能

条件	調査用データ (11 話)						評価用データ (4 話)					
	正解	脱落	挿入	Recall	Precision	F 尺度	正解	脱落	挿入	Recall	Precision	F 尺度
場	76	91	158	45.5	32.5	37.9	27	24	47	52.9	36.5	43.2
時	78	89	92	46.7	45.9	46.3	24	27	28	47.0	46.2	46.6
人	96	71	180	57.5	34.8	43.3	29	22	39	56.9	42.6	48.7
場&時	83	84	105	49.7	44.1	46.8	25	26	33	49.0	43.1	45.9
場&人	102	65	170	61.1	37.5	46.5	33	18	56	64.7	37.1	47.1
時&人	111	56	143	66.5	43.7	52.7	35	16	44	68.6	44.3	53.8
時&人 (2)	117	50	151	70.1	43.6	53.8	38	13	47	74.5	44.7	55.9
場&時&人	101	66	107	60.5	48.6	53.9	31	20	35	60.8	47.0	53.0
改良	101	66	103	60.5	49.5	54.4	29	22	34	56.9	46.0	50.9
改良 (2)	115	52	146	68.9	44.1	53.7	38	13	46	74.5	45.2	56.3

今回、場所等の候補となる語句の異なり数に基づいてシーン境界を推定する方法を試みた。これは、新しい文を処理対象とした場合に、前の文に現われていない場所等を示す語句が現われても、それだけではシーン境界と判断することはできないためである。そのような条件を念頭に、一つの方法として試みた。しかし、この手法では、今後の研究に必要と思われる性能は得られなかった。

十分な性能が得られなかった理由の一つとして、挿入誤りが多いことが挙げられる。今後、挿入誤りを減らすことでの性能向上が検討課題である。

なお、本実験では、表 2 の d の項目にあたる条件は、シーン境界推定の基準に明示的には取り入れていないが、調査用データでは d にあたる箇所 30 箇所中 20 箇所 (67%) が、評価用データでは 8 箇所中 5 箇所 (62%) が検出されている。

4 関連研究

テキストを分割する代表的な手法としては、Hearst による Text Tiling 法が挙げられる [14]。Text Tiling 法では窓内における語彙の重なり尺度に基づく類似度

より、分割位置を推定する。だが、物語の場合、場所や時間に関する語彙が提示されるのはそれらに変化があった場合の一度のみであること、登場人物の人数は限られていること、あるいは記述が簡潔であるという特徴がある。そのため、Text Tiling 法は今回の物語のシーン分割に適しているとは言えないと判断した。

物語を対象としたテキストの分割に関しては、今らによる研究がある [15]。今らは、物語テキストの段落番号、文番号、場所に付与された番号を用い、クラスタリングによりシーン分割を行なっている。システムの出力を人間が評価した結果、シーン区切りの 90% が妥当であるという評価を得ている。

福田らは、話し言葉の書き起こしテキストの話題分割において、F 尺度で 45 という結果を得ている [12]。

柴田らは、料理教示発話を対象として話題分割を行ない、F 尺度で 72 という結果を得ている [13]。

加藤らは、レストラン検索対話を対象とし、対話の構造を求める手法について報告している [16]。

馬場らは、WWW 上に公開されている小説から、人名辞典を用いて、人物抽出を行なっている [17]。

三宅らは、BMCL 法を提案し、「星の王子さま」のストーリー・マップを表わす意味ネットワークを構築し

ている [18]。

物語の生成に関しては、筆者は、手続き的なプロット作成ツール “scena1” を公開している [19]。

藤田らは、登場人物に着目した物語生成を検討している [20]。

また、佐久間らは、プロットのファンクションを利用した、物語作成支援システムを検討している [21]。

5 まとめ

本研究では、計算機による物語理解のために物語を「分節」化すること、および構造主義における物語の構造分析により得られた知見を工学的に利用するための研究の端緒として、場所、時間、登場人物に基づいたシーン境界推定手法を予備実験として試みた。結果として、場所、時間、登場人物についての語句の文間での異なり数に応じたペナルティに基づく方法により、F 尺度で、調査用データで 54、評価用データで 56 という結果が得られた。

今後、本研究の結果をベースラインとし、場所候補、時間候補、登場人物候補の語彙ないし概念によりペナルティを変える、あるいは文としての意味および文間の関係を考慮に入れる等の方法によってシーン境界推定の性能の改善を試みる。また、得られたシーン等と「プロットのファンクション」や登場人物等の対応付けを行なうことを前提に、シーン等の内容の要約ないしは抽象化に関する研究を行なう。

参考文献

- [1] Roland Barthes, “Introduction à l’analyse structurale des récits”, *Communication*, 8, novembre (1966). [ロラン・バルト, 「物語の構造分析」, 花輪光 訳, みすず書房 (1979).]
- [2] 小林 聡, 山口 優, 中川 聖一, “表層的言語情報と韻律情報を用いた講演音声の重要文抽出による要約”, *自然言語処理*, vol. 12, No. 6. pp.3-24 (2005 年).
- [3] Владимир Я. Пропп, “Морфология сказки”, Изд. 2е, Москва, Наука (1969). [ウラジーミル・R・プロップ, 「昔話の形態学」, 北岡 誠司・福田 美智代 訳, 水声社 (1987).]
- [4] A. J. Greimas, “Sémantique structurale”, Larousse (1966). [A. J. グレマス, 「構造意味論」, 田島 宏・鳥居 正文 訳, 紀伊国屋書店 (1988).]
- [5] Robert Stam, Robert Burgoyne, Sandy Flitterman-Lewis, “New Vocabularies in Film Semiotics”, England, Routledge (1992). [ロバート・スタム, ロバート・バーゴイン, サンディ・フリッターマン=ルイス, 「映画記号論入門」, 丸山 修, エグリントンみか, 深谷 公宣, 森野 聡子 訳, 松柏社 (2006).]
- [6] 川内 彩友美, “決定版 まんが日本昔ばなし 101”, 講談社 (1997).
- [7] “日本語形態素解析システム JUMAN”, <http://nlp.kuee.kyoto-u.ac.jp/nl-resource/juman.html>
- [8] “日本語構文解析システム KNP”, <http://nlp.kuee.kyoto-u.ac.jp/nl-resource/knp.html>
- [9] NTT コミュニケーション科学基礎研究所, “日本語語彙体系”, 岩波書店 (1999).
- [10] “EB ライブラリ”, <http://www.sra.co.jp/people/m-kasahr/eb/index.html>
- [11] “RubyEB”, <http://rubyeb.sourceforge.net/index-ja.html>
- [12] 福田 雅志, 延澤 志保, 太原 育夫, “語彙的結束性に基づく話し言葉のテキストセグメンテーション”, *言語処理学会 第 11 回 年次大会予稿集*, pp.620-623 (2005).
- [13] 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”, *言語処理学会 第 11 回 年次大会予稿集*, pp. 109-112 (2005).
- [14] M. A. Hearst, “Text Tiling: Segmenting Text into Multiparagraph Subtopic Passages”, *Computational Linguistics*, vol. 23, no. 1, pp. 33-64 (1997).
- [15] 今 誠一, 吉田 文彦, 内田 理, 菊池 浩明, 中西 祥八郎, “昔話の自動シナリオ化システムの構築”, *言語処理学会 第 11 回 年次大会予稿集*, pp. 317-320 (2005).
- [16] 加藤 真吾, 松原 茂樹, 山口 由紀子, 河口 信夫, “発話意図タグに基づく対話構造の記述”, *言語処理学会 第 11 回 年次大会予稿集*, pp. 129-132 (2005).
- [17] 馬場 こづえ, 藤井 敦, 石川 徹也, “小説テキストを対象としたジャンル推定と人物抽出”, *言語処理学会 第 11 回 年次大会予稿集*, pp. 157-160 (2005).
- [18] 三宅 真紀, Jaeyoung Jung, 赤間 啓之, “グラフクラスタリングとパターン分類を併用したストーリー・マップ生成の試み”, *言語処理学会 第 12 回 年次大会予稿集*, pp.644-647 (2006).
- [19] 小林 聡, “scena1”, <http://skoba.hp.infoseek.co.jp/scena/scena1/scena1.html> (1995).
- [20] 藤田 智之, 水嶋 玲子, 小川 均, “登場人物の意思に基づいた物語生成”, *情報処理学会 自然言語処理研究報告*, 1997-NL-122, pp.145-152(1997).
- [21] 佐久間 友子, 小方 孝, “物語自動生成によるストーリー作成支援の検討”, *人工知能学会第二種研究会 ことば工学研究会 (第 24 回) 資料*, pp. 83-88 (2006).