

新近性効果の減衰曲線を加味した顕現性計算手法に基づく 話題遷移の可視化

白松 俊 駒谷 和範 尾形 哲也 奥乃 博

京都大学大学院情報学研究科

{siramatu,komatani,ogata,okuno}@kuis.kyoto-u.ac.jp

1 はじめに

背景 本研究では、長い談話 (e.g., 議事録, 裁判記録) の流れをユーザが容易に把握できる談話閲覧インタフェースの実現を目指す。動的に変化する流れの把握を容易にするには、談話参加者の注目対象の遷移をどのように可視化するかが重要な課題である。そのためにわれわれは、参加者達が各単語に注目する度合いを定量化し、その遷移を視覚化する。この注目度を、本稿では顕現性 (*salience*) と呼ぶ。顕現性は、発話ごとに推移する心理量である。

仮定 本来、各参加者にとっての顕現性は主観に依存するが、本稿では客観的に観測可能な顕現性を対象とする。すなわち、参加者間で顕現性の分布が共有されている協調的な状況を仮定する。これは、談話の流れの可視化には客観的な顕現性が必要なためである。

具体例 目指す話題遷移図の具体例 (約 830 文から成る議事録) を図 1 に示す。横軸が発話系列、縦軸が顕現性の尺度を表す。われわれは、このようなグラフを *SalienceGraph* と名付ける。これにより、発話系列中のどこで所望の議論が行われているかをユーザが把握し易くすることで、Shneiderman の *Visual-Information Seeking Mantra*, すなわち “Overview first, zoom and filter, then details on demand” [3] を満たす談話インタフェースが実現可能と考える。

技術的要求 *SalienceGraph* の縦軸にあたる顕現性の計算手法は、以下の要件を満たす必要がある。

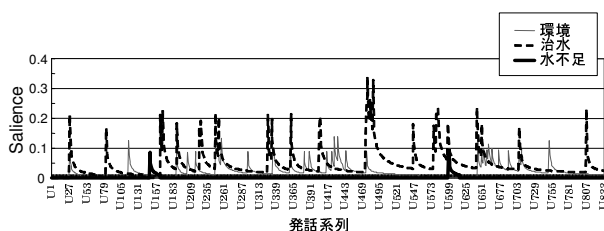


図 1 *SalienceGraph*: 顕現性に基づく談話遷移の可視化

- 時系列的遷移: 顕現性の動的な変化を表現するために、時系列的遷移を扱うことが必要不可欠である。
- 定量化: 工学的に扱い易いように、(順序尺度ではなく) 定量的な尺度として定義する必要がある。
- 素性の統合基準: 多様な顕現性の影響要因 (e.g., 単語頻度, 助詞, etc.) を統合する基準が必要である。
- 評価尺度: 計算手法の最適化のために、顕現性計算のパフォーマンスを測る評価尺度が必要である。

本稿で扱う課題 われわれは過去の研究で、顕現性の尺度として参照確率を定義した [4]。これは、「目立っている単語ほど継続的に参照されやすい」という性質を利用した尺度であり、センタリング理論 [1] との整合性が実データ上で確認されている。発話 U_i における実体 e の顕現性は、以下のように定義される。

$$\begin{aligned} (U_i \text{ での } e \text{ の顕現性}) &= \Pr(\exists w \xrightarrow{\text{ref}} e \text{ in } U_{i+1} | \text{pre}(U_i)) \\ &= \Pr(e | \text{pre}(U_i)) \end{aligned}$$

ただし、 U_{i+1} は後続発話を、 $\text{pre}(U_i)$ は先行文脈 $[U_1, \dots, U_i]$ を、 $w \xrightarrow{\text{ref}} e$ は単語 w が実体 e を参照することを表す。 $\Pr(e | \text{pre}(U_i))$ は参照確率の略記法である。計算は、ロジスティック回帰モデルで複数の素性を統合することで行う。参照確率は定量化されており、複数の素性の統合基準を備えているが、以下の課題を残していた。

- (1) 「最近の事象ほど思い出され易い」という新近性効果 (*recency effect*) [2] の減衰曲線を直接扱っていなかったため、時系列的遷移の扱いに改良の余地が残されていた。
- (2) 最適化のための評価尺度が明らかでなかった。

2 新近性効果の減衰曲線を表す窓関数の導入

新近性効果とは、図 2 に示す系列位置曲線 (*serial position curve*) のうち、右側の楕円で囲まれた部分

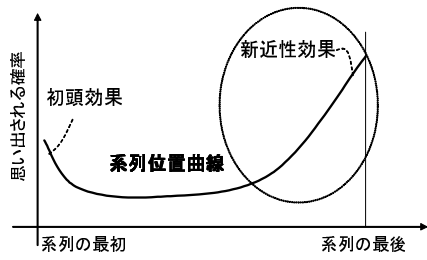


図 2 系列位置曲線と新近性効果

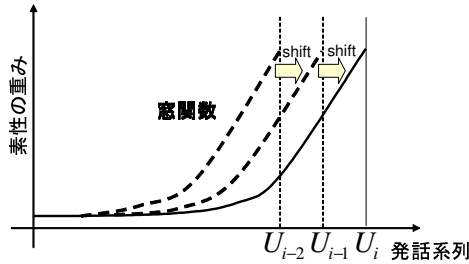


図 3 1 発話毎にシフトする窓関数で新近性効果を表現

のことである。この曲線は、被験者に孤立単語の系列を提示し、最後に思い出してもらおうという実験によって得られる。ただし、談話文脈における新近性効果の性質は未解明である。われわれは、より正確に顕現性の動的遷移を表すために、1 発話毎にシフトする窓関数 (*window function*) によって新近性効果の減衰曲線を表現する (図 3)。図 4 に示すように、窓関数を掛けることで過去の出現箇所ほど素性の重みを小さくすれば、談話文脈における新近性効果を扱うことができる。ただし、図中の $W(\text{dist})$ は $W(\text{dist}(w, U_{i+1}))$ の略記であり、 $\text{dist}(w, U_{i+1})$ は、 $w \xrightarrow{\text{ref}} e$ in $\text{pre}(U_i)$ たる単語 w から後続発話 U_{i+1} への発話距離を表す。

具体的には、以下の 4 種の窓関数について後述する評価尺度が最大となるパラメタを求める。そして、最適な窓関数を 1 つ選択する。

- 矩形窓 (パラメタ: k)

$$W(\text{dist}) = \begin{cases} 1 & \text{dist} \leq k \\ 0 & \text{otherwise} \end{cases}$$

- ガウス窓 (パラメタ: σ)

$$W(\text{dist}) = \exp\left(-\frac{\text{dist}^2}{\sigma^2}\right)$$

- 指数窓 (パラメタ: T)

$$W(\text{dist}) = \exp\left(-\frac{\text{dist}}{T}\right)$$

- 逆数窓 (パラメタ: d)

$$W(\text{dist}) = \frac{1}{\text{dist}^d}$$

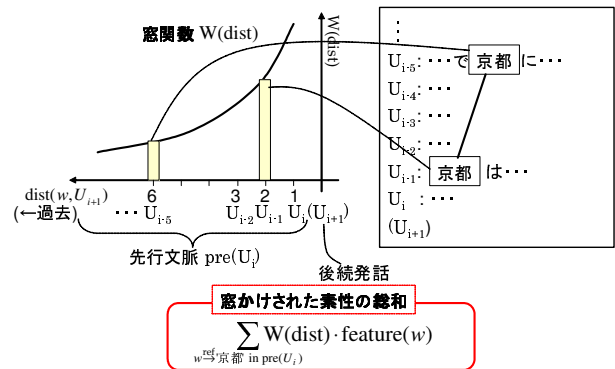


図 4 窓関数を用いた素性の重み付け

3 評価尺度: 後続発話における参照の予測性能

ある顕現性の推定手法 m の評価尺度 $\text{evalSal}(m)$ を、以下のように定義する。

$$\text{evalSal}(m) = \text{cor}\left([\text{sal}_m(e|\text{pre}(U_i))]_{\langle e, U_i \rangle}, [\text{isRef}(e, U_{i+1})]_{\langle e, U_i \rangle}\right)$$

ただし、 $\text{cor}(x, y)$ は x と y の相関係数、 $\langle e, U_i \rangle$ は実体 e と発話 U_i の組から成るサンプル、 $\text{sal}_m(e|\text{pre}(U_i))$ は m により計算される U_i における e の顕現性の値、 $\text{isRef}(e, U_{i+1})$ は後続発話における e への参照の正解

$$\text{isRef}(e, U_{i+1}) = \begin{cases} 1 & (\exists w \xrightarrow{\text{ref}} e \text{ in } U_{i+1}) \\ 0 & (\text{otherwise}) \end{cases}$$

である。この $\text{evalSal}(m)$ は、テストセットコーパスを用いて計測され、後続発話における参照の予測性能を表す。つまり、「後続発話での参照の予測性能が高い尺度ほど、良い顕現性尺度である」という仮定に基づき、顕現性計算手法を最適化することができる。

4 実験: 計算手法の最適化

コーパス 実験には以下の 2 種のコーパスを用いる。

- 話し言葉: 日本語話し言葉コーパス (CSJ) に含まれる 4 つの 2 者自由対話 (インタビュー D03F0001, D03M0013, D03F0040, D03M0048)。1,780 発話単位 (IPUs; inter-pause units), 1 発話単位あたり 6.92 形態素を含む。
- 新聞記事: 1994 年の毎日新聞 3,000 記事。63,221 発話単位 (述語節), 1 発話単位あたり 10.79 形態素を含む。

表 1 各素性に最適化された窓関数の比較 (CSJ)

素性	窓関数	最適なパラメタ	evalSal(m)
出現頻度	矩形窓	$k = 9$	0.1048
	ガウス窓	$\sigma = 3.66$	0.1667
	指数窓	$T = 2.51$	0.1727
	逆数窓	$d = 1.35$	0.1731
文法役割	矩形窓	$k = 9$	0.1195
	ガウス窓	$\sigma = 5.01$	0.1928
	指数窓	$T = 3.39$	0.2013
	逆数窓	$d = 1.14$	0.2063
品詞	矩形窓	$k = 9$	0.1200
	ガウス窓	$\sigma = 2.52$	0.2226
	指数窓	$T = 2.19$	0.2308
	逆数窓	$d = 1.24$	0.2390

表 2 各素性に最適化された窓関数の比較 (毎日新聞)

素性	窓関数	最適なパラメタ	evalSal(m)
出現頻度	矩形窓	$k = 1$	0.3013
	ガウス窓	$\sigma = 0.99$	0.3467
	指数窓	$T = 0.33$	0.3467
	逆数窓	$d = 4.27$	0.3468
文法役割	矩形窓	$k = 1$	0.2991
	ガウス窓	$\sigma = 1.14$	0.3682
	指数窓	$T = 0.45$	0.3685
	逆数窓	$d = 3.01$	0.3696
品詞	矩形窓	$k = 1$	0.2985
	ガウス窓	$\sigma = 1.04$	0.3537
	指数窓	$T = 0.37$	0.3537
	逆数窓	$d = 3.80$	0.3540

それぞれ、統語情報と照応情報を表す GDA (Global Document Annotation) タグが付与されている。統語タグは半自動で、照応タグは人手で付与した。

最適な窓関数の決定 参照確率はロジスティック回帰によって計算される [4] が、まずは回帰を用いず、各素性に重み付けする窓関数の部分最適化を行った。出現頻度、文法役割、品詞のそれぞれに対して窓をかけて総和した値 (図 4) を顕現性推定値と見なした場合に、評価尺度 evalSal(m) を最大化する最適な窓関数を求めた。その結果を表 1, 2 に示す。これにより、談話文脈における新近性効果の減衰曲線を表す窓関数としては、逆数窓が最適であることが示された。更に、CSJ で最適な窓関数は、毎日新聞で最適な窓関数よりもなだらかであった。新聞記事は紙面の制約によって注目対象が素早く遷移するのにに対し、対象とした CSJ のインタビューはそのような制約がほとんど

表 3 素性の候補

素性	説明
gram(w)	$w \xrightarrow{\text{ref}} e$ が係っている機能語 (文法役割)
pos(w)	$w \xrightarrow{\text{ref}} e$ の品詞
title(e)	e がタイトル中で参照されているか
dist(w, U_{i+1})	U_j 中で $w \xrightarrow{\text{ref}} e$ のとき, $(i+1) - j$
freq((e, U_i))	$\frac{1}{i} (\# w \xrightarrow{\text{ref}} e \text{ in pre}(U_i))$

表 4 各素性単独での評価尺度の比較 (CSJ)

gram	pos	dist	freq	回帰	evalSal(m)
✓					0.0475
	✓				0.0526
		✓			0.1519
			✓		0.0437

表 5 各素性単独での評価尺度の比較 (毎日新聞)

gram	pos	title	dist	freq	回帰	evalSal(m)
✓						0.1171
	✓					0.0817
		✓				0.0644
			✓			0.2831
				✓		0.2519

無いため、相対的に新近性効果の減衰曲線がなだらかであると考えられる。

素性の影響力の調査 表 3 に示す素性の候補に対し、各素性単独での評価尺度 evalSal(m) を計測した結果を表 4, 5 に示す。これにより、話し言葉、新聞記事の双方でともに dist, すなわち新近性効果が最も影響力の強い素性であることが示された。

最適な素性セットの選択 参照確率 $\text{Pr}(e|\text{pre}(U_i))$ を計算するためのロジスティック回帰で用いる素性セットの最適化を行った結果、両コーパスで図 5 の計算手法が最適であった。ただし、図の avgPr は離散的な素性に値を割り当てる関数である。単語 w の離散素性 $\text{feature}(w) = x$ に対する avgPr は、以下のように定義される。

$$\text{avgPr}(x) = \frac{\#\{w; C(x, w, e, U_i) \wedge \exists w' \xrightarrow{\text{ref}} e \text{ in } U_{i+1}\}}{\#\{w; C(x, w, e, U_i)\}}$$

$$C(x, w, e, U_i) \equiv (w \xrightarrow{\text{ref}} e \text{ in pre}(U_i) \wedge \text{feature}(w) = x)$$

図 5 が最適となった根拠が、表 6, 7 である。これらは、素性セットの組み合わせを変え、10 分割交差検定

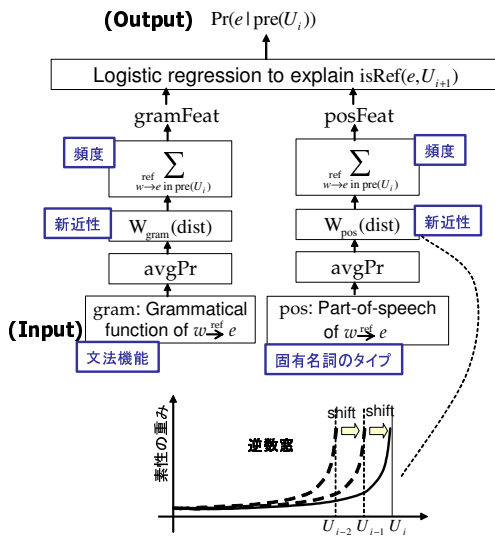


図5 評価尺度最大となった参照確率の計算手法

表6 素性セットの選択 (CSJ)

$gram$	pos	$W(dist)$	$W_{ref}^{w \rightarrow e}$ in $pre(U_i)$	回帰	$evalSal(m)$
✓		✓	✓	✓	0.3458
	✓	✓	✓	✓	0.3591
✓	✓	✓	✓	✓	0.3652

表7 素性セットの選択 (毎日新聞)

$gram$	pos	$title$	$W(dist)$	$W_{ref}^{w \rightarrow e}$ in $pre(U_i)$	回帰	$evalSal(m)$
✓			✓	✓	✓	0.3687
	✓		✓	✓	✓	0.3437
✓	✓		✓	✓	✓	0.3680
		✓	✓	✓	✓	0.0860
✓		✓	✓	✓	✓	0.3638
	✓	✓	✓	✓	✓	0.3395
✓	✓	✓	✓	✓	✓	0.3631

によって評価尺度 $evalSal(m)$ を計測した結果である。評価尺度 $evalSal(m)$ は、CSJ で 0.3652、毎日新聞で 0.3680 となった。

矩形窓 TF との性能比較 本手法を、TF(単語の頻度)に矩形窓をかけた単純な手法と比較する。前述の表1, 2の最初の行(矩形窓をかけた出現頻度)が、すなわち TF に矩形窓をかけたものである。参照の予測

性能を表す $evalSal(m)$ が、CSJ で矩形窓 TF: 0.1048 → 本手法: 0.3652、毎日新聞で矩形窓 TF: 0.3013 → 本手法: 0.3680 となった。よって、特に話し言葉で、最適な窓関数と素性セットの選択が有効であることが明らかになった。

得られた知見 上記の実験で得られた知見を示す。

- (1) 談話文脈における新近性効果の減衰曲線を扱うには、逆数窓が適していた。また、話し言葉は新聞記事に比べて新近性効果の減衰がなだらかであった。
- (2) 評価関数 $evalSal(m)$ に基づく最適な窓関数と素性セットの選択は、特に話し言葉で有効であった。

5 関連研究

活性化拡散モデル (*Spreading activation model*) や PLSA (*Probabilistic Latent Semantic Analysis*) 等も顕現性計算に利用できる可能性がある。今後は、矩形窓 TF だけでなく、これらとの比較実験も行う。

6 まとめ

図1に示した SaliencyGraph で談話の流れを可視化するため、顕現性を表す参照確率の計算手法に2つの拡張(新近性効果の減衰曲線を表す窓関数と、後続発話での参照の予測性能を表す評価尺度)を加えた。これにより、文脈の動的変化をより良く捉えるよう計算手法を最適化できた。今後は、SaliencyGraph を用いた談話閲覧システムを開発する予定である。

謝辞 新聞記事 GDA コーパスの研究利用を許諾された三菱電機株式会社と調整して下さった GSK に感謝する。本研究は科研費補助金(特別研究員奨励費)の支援を受けた。

参考文献

- [1] B. Grosz, A. Joshi, and S. Weinstein. Centering: A Framework for Modeling the Local Coherence of Discourse. *Computational Linguistics*, Vol. 21, No. 2, pp. 203–226, June 1995.
- [2] B.B. Murdock. The Serial Position Effect in Free Recall. *Journal of Experimental Psychology*, Vol. 64, pp. 482–488, 1962.
- [3] B. Shneiderman and C. Plaisant. *Designing the User Interface: Strategies for Effective Human-Computer Interaction (4th Edition)*. Addison-Wesley, 2004.
- [4] 白松俊, 駒谷和範, 橋田浩一, 尾形哲也, 奥乃博. ゲーム理論に基づく参照結束性のモデル化と日本語・英語の大規模コーパスを用いた統計的検証. *自然言語処理*, Vol. 14, No. 5, pp. 199–239, 2007.