

動画コンテンツに対する コメントの自己組織化と 要約・概念学習への応用

安田女子短期大学
坂口 琢哉

はじめに

▶ 動画コメントとは

- ユーザが視聴中の動画やTV番組等に対して、リアルタイムに付けるコメント
- Webサービスの例: Twitter、2ch掲示板、ニコニコ動画等
- 例: 「〇〇は可愛いなあ」「犯人は〇〇だろ」「キターー」など
- 要素: コメント本文、メタデータ(投稿日時、ユーザID等)
- 短文、直観的、短絡的な傾向

▶ 研究の目的と特徴

- 大量の動画コメントを、書込日時や類似度によって収束
- 大量で持続的に生産されるデータの活用
- 動画要約・概念学習等への応用の可能性

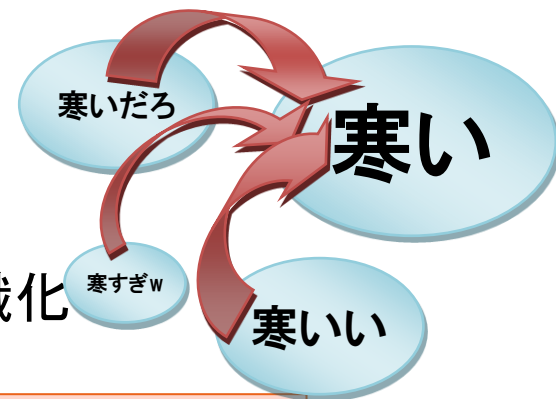
提案手法

- ▶ 取得したコメント間の類似度を計算
 - 一定時間枠(±60sec / ±180sec / ±600sec)内に存在する任意のコメント M , N を対象
 - **3-Gram法**に基づき類似度 S_{MN} を計算

$$S_{MN} = \frac{\sum_{i=1}^{m-2} f(M_i M_{i+1} M_{i+2}, N)}{(m-2) + (n-2) - \sum_{i=1}^{m-2} f(M_i M_{i+1} M_{i+2}, N)}$$


ただし、 $f(A, B) = \begin{cases} 1(A \in B) \\ 0(A \notin B) \end{cases}$

- ▶ 活性値の変化による自己組織化
 - コメント M に対する活性値 a_M (=重要度)
 - **類似のコメント間で活性値をやりとり**
 - 値は、**低い方から高い方へ**と移動 → 自己組織化
 - 値の移動量は、類似度 S_{MN} に依存



$$\begin{cases} a_M + = da_N S_{MN} \\ a_N - = da_N S_{MN} \end{cases} \quad \text{ただし、} a_M \geq a_N, \quad d \text{ は変化量を表す係数}$$

実験結果

	1	2	3	4	5	6	7	8	9	計
	0	0	0	0	0	0	0	0	0	0
	0	0	0	1	2	1	0	0	X	4

▶ 実験概要

- 対象動画: プロ野球広島戦(2011/9/4)の試合
- 対象コメント: 2ch実況板の広島応援スレッド(計9673レス)
- 実験時のパラメタ: $d=0.1$ 、 $\alpha=0.1$ 、 $k=10$

▶ 結果

- 試合の重要局面において、ピンポイントでコメントが活性化
- 活性化されたコメントから、ある程度場面の状況を推測可能

▶ 今後の展望

- コメントを加工し、タグとして最適化
 - 動画要約への応用
- コメントの多様性を確保
 - 概念学習への応用

