

MuST

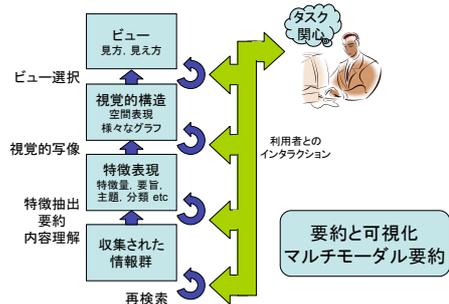
Multimodal Summarization for Trend Information

Organizers:
 加藤恒昭 (東京大学)
 松下光範 (NTT CS研)

位置づけ

- 対話的かつ探索的な情報活用の支援
 - 収集された情報の全体像の概観
 - 関心の絞り込み・具体化や変更
 - 必要な詳細情報へのアクセス
- 言語情報と非言語情報の活用
 - 言語情報と非言語情報の横断的利用
 - マルチモーダルプレゼンテーション

対話的かつ探索的な情報活用の支援



動向情報への関心

- 利用者の関心への最初の回答としての動向
 - 「06年からのゲーム機業界はどんなでしたか？」
 - 「最近、ガソリンの価格はどう動いている？」
 - 「去年の台風はひどかったですか？」
- 様々な興味深い性質
 - 一定期間にわたる情報の総合的まとめあげ
 - 時間、地理的空間に止まらない様々な軸
 - 事実に加えて、解釈、原因や影響の推測、予測
 - テキスト情報+数値情報

MuST:動向情報の要約と可視化に関するワークショップ

- 2005年初めより、NTCIR Pilot Taskとして約1年サイクルで実施
- 共通の関心に基づき、共通の研究資源を用いた緩い意味での共通課題への取り組み
 - コミュニティの形成
 - ツールやコーパス類の蓄積
- 2回の成果報告、NTCIR Workshop Meetingでも一部発表

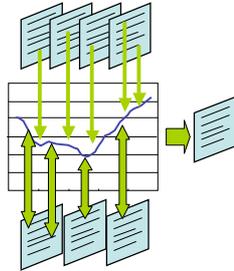
→ 共通的な課題・関心が徐々に明らかに

タスクの概要

- 情報活用を目的に、動向情報を中心としたマルチモーダル要約についての研究を加速し、その検討・評価を行う
- 関心・目的の具体化に応じたふたつのアプローチをとる
 - 評価課題: 具体化された研究課題について、定量的な評価を検討し、研究を進める
 - 自由課題: 萌芽的な研究課題について、模索的、提案的な研究を進める

評価課題

- テキストからの時系列統計量の抽出と可視化
- 時系列数値情報の言語化
- 時系列数値情報とテキスト情報とのアラインメント



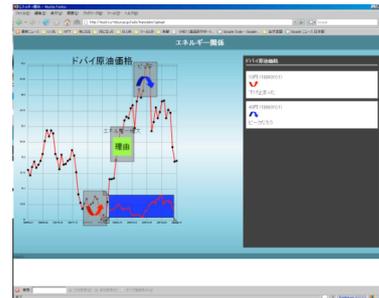
資源

- MuST データセット
 - 1998,1999年の毎日新聞記事より27トピック581記事について注釈付けを行ったもの
 - 毎日新聞1998年,1999年(+2000年)全記事
- 可視化プラットフォーム
 - 情報活用を目的とした可視化システム構築のためのプラットフォーム
 - オープンソースで提供予定
- これまでに蓄積されたデータ等

MuSTデータセット注釈例

```
<unit stat="レギュラーガソリンの全国平均店頭価格"><del type="src">石油情報センターが16日発表した給油所石油製品市況調査によると、</del><date gra="週" abs="19990617">今週</date>調査の<name part="head">ガソリン価格(レギュラー1リットル)</name>は<name part="foot">全国平均</name>で<val>92円</val>となり、<date gra="週" abs="19990610">前週</date>の<name part="foot">平均</name>に比べ<rel type="diff">1円</rel>上昇した</unit>。<unit stat="レギュラーガソリンの全国平均店頭価格"><name part="foot">全国平均</name>の上昇は<date gra="旬" abs="19990201">2月上旬</date>以来、<dur gra="月">4か月</dur>ぶりだ</unit>。
```

可視化プラットフォーム



スケジュール

- 07.11末 参加申し込み締切
- 07.12 第一回ラウンドテーブルミーティング
評価課題の具体化
- 08.2 評価課題予備実験(Dry Run)
- 08.3 中間報告会
- 08.6 評価課題本実験(Formal Run)
- 08.8 第二回ラウンドテーブルミーティング
評価課題の評価に関する議論
- 09.9末 NTCIR7 レポート提出

連絡先等

- Organizers
 - 加藤恒昭 kato@boz.c.u-tokyo.ac.jp
 - 松下光範 mat@cslab.kecl.ntt.co.jp
- Wiki page
 - 一般用 <http://must.c.u-tokyo.ac.jp>
 - 参加者用 <http://.../participants/wiki>
- Mailing List
 - 一般用 must@cslab.kecl.ntt.co.jp
 - 参加者用 must-participants@...