Towards Multi-modal Extraction and Summarization of Conversations

Raymond T. Ng

(University of British Columbia, Canada)

(www.cs.ubc.ca/~rng)

January 2010

Acknowledgements

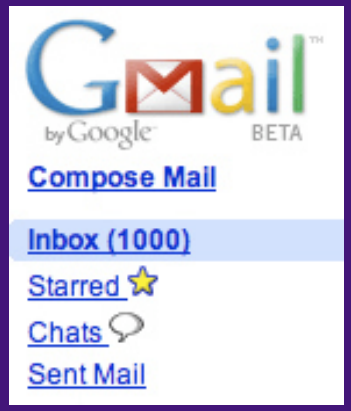
Giuseppe Carenini (faculty colleague, www.cs.ubc.ca/~carenini)
Gabriel Murray (postdoctoral fellow)
Jackie Cheung
Nicholas Fitzgerald
Shafiq Joty
Adam Pauls
Lucas Rizoli
Jan Ulrich
Xiaodong Zhou
Ed Zwart

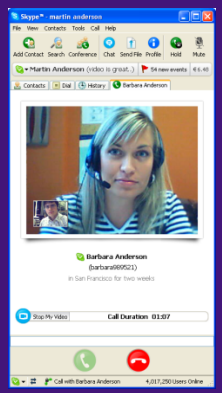
Text Data Explosion

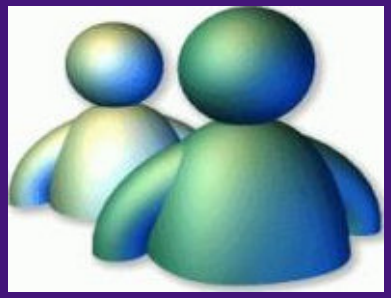
- Past Decade witnessed an exponential growth in text data accumulation
  - Formal Documents: newspapers, reports, webpages, etc.
  - Informal Documents: emails, blogs, MSN, user reviews, etc.
- Text data management and data mining are important areas for research and development
- This talk focuses on *conversational* data
  - But there are many other text data management problems, e.g., data cleansing, duplicate identification, etc.

Motivations

- In our daily lives, we have conversations with people in many different modalities
  - Emails, meetings, telephone, videoconferencing, instant messaging, blogs, forums, etc.
  - The Web has significantly increased the volume and the complexity







Conversation Summarization

- With the high volume of conversational data being generated every day, summarization can be beneficial by creating concise overviews that aid quick access to the data
  - Expediting personal browsing of the data
  - Helping preserve corporate memory
  - Content linking
- This talk will consider both *extractive* and *abstractive* summaries

Business Intelligence Scenario

- In a meeting, the VP on marketing raised the topic of developing a new product
- Subsequent emails continued the discussion on:
  - Web documents describing similar products;
  - User reviews on those products
- This conversation spans meeting notes, emails, web documents, customer reviews and blogs
- How to automate the generation of a concise summary/report of the (ongoing) conversation?
  - How were the decisions made? What were the issues? Etc.

Key Research Questions

1. How to recognize that (text) data from multiple modalities are for the same conversation?
2. How to summarize the text:
   - From a single modality, e.g., meeting notes, emails, customer reviews, blogs?
   - From across multiple modalities/genres? Is there a universal set of features, or are the features genre-specific?

Outline of the Rest of the Talk

How to recognize that (text) data from multiple modalities are for the same conversation?

- Our work on forming email conversations

How to summarize the text:

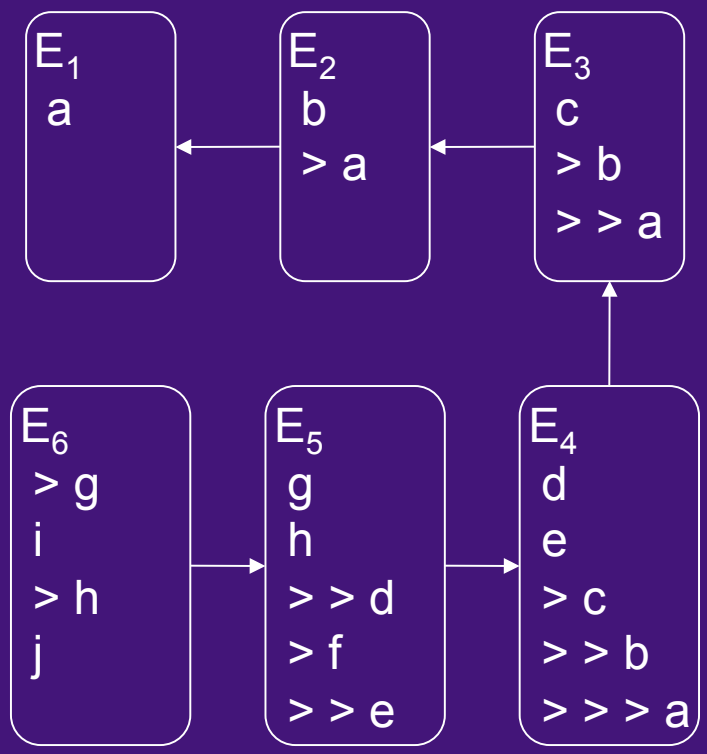
- From a single modality, e.g., meeting notes, emails, customer reviews, blogs?
- Our work on summarizing emails, meeting notes and customer reviews
- From across multiple modalities/genres? Is there a universal set of features, or are the features genre-specific?
- Our experience on meeting + email summarization

Email Conversation Extraction and Summarization

Extracting Conversation from Quotations

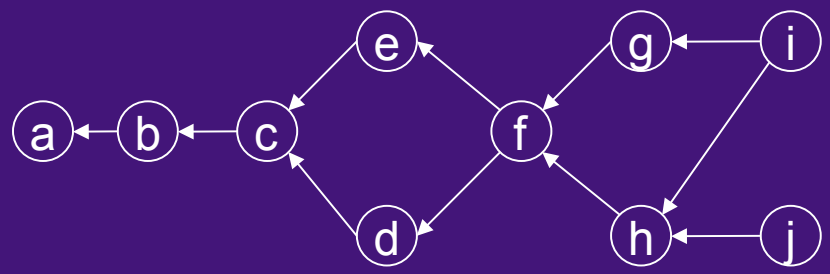
- Email header and threading useful but inaccurate and crude
  - Using the most recent, but wrong, header
  - Multiple topics in one email
- We analyzed the actual body of the emails, particularly the quotations
  - Selective quotations may indicate the intention of the sender in responding to specific parts of the email
  - i.e., from email thread (node = email) to Fragment Quotation Graph (node = text fragment in email)

Fragment Quotation Graph



An email conversation with 6 emails.

- nodes
  - Identify quotations and new fragments
- edges
  - Neighbouring quotations



Extractive Summary of an Email Conversation

- Multi-document summarization a well-established field (e.g., Radev's 2004 SIGIR tutorial)
- Developed mainly for more formal documents, e.g., news, reports
- Emails is a different genre: less grammatical, can be much shorter, less focused
- Our objective is to find an extractive summary: select the top-k sentences from the emails that best represent the content

Summarization by Clue Words

- A *clue word* in a node is a non-stop word that also appears (modulo stemming) in a parent or a child node in the quotation graph
- i.e., a sentence may be more important if it contains a word that is reused in a reply
- Sentences are ranked by their clue word score
- We showed that this provides a better summary than one generated by MEAD, a state-of-the art multi-document summarizer

Supervised Email Summarization

Clue word summarization is *unsupervised*

Rambow et al. used a *supervised* approach

- Text features, e.g., centroid similarity, tf-idf sum and average, length, etc.
- Email features, e.g., relative position, number of replies and recipients, subject similarity

However, the two approaches are not incompatible as clue words can be incorporated as a feature in classifier training

An Enhanced Supervised Framework

Apart from clue words, we added other features: speech acts (Carvalho et al.), meta sentences (Murray et al.) and subjectivity (Carenini et al.)

E.g., Propose, Request, Commit, and Meeting:

A Propose sentence proposes a joint activity;

a Request sentence asks the recipient to perform an activity;

a Commit sentence commits the sender to some future course of action; and

a Meeting sentence is regarding a joint activity in time/space

All these new features combined to give a better summary

Meeting Summarization

Meeting Summarization

- Meeting summarization relatively underdeveloped
- A sensible place to begin is to apply text techniques
- But meeting can be disfluent, spontaneous – i.e. difficult
- Is there anything to be gained by treating meeting notes as not only text, i.e., genre-specific features?

Results Overview

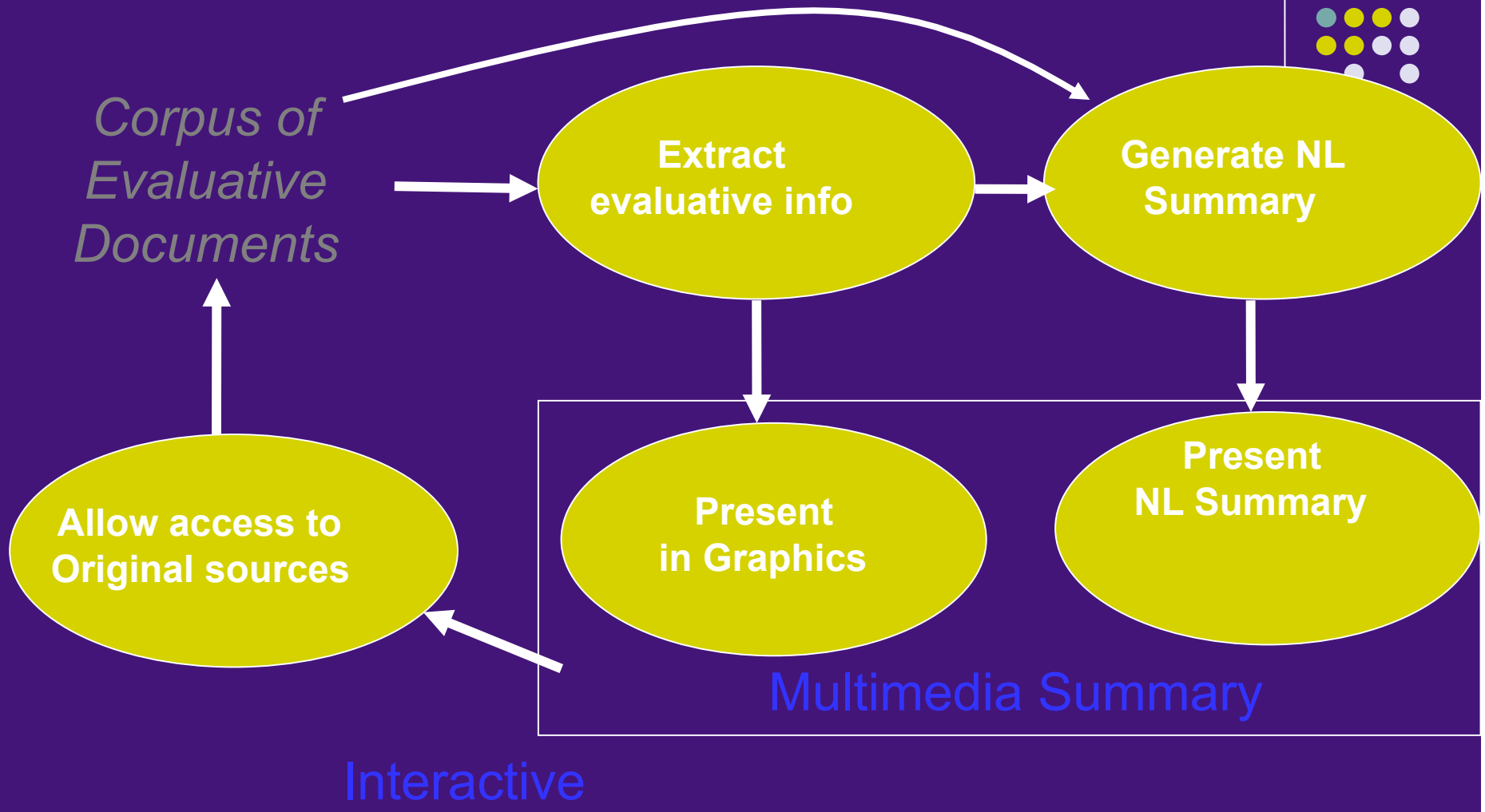
- We used human annotators to identify important sentences: gold standard
- We built classifiers based on different types of features and compared against the gold standard
- Features that are effective
  - Structural Features: turn position
  - Speaker features: dominance (in length, dialogue acts)
  - Speech features: loudness, rate-of-speech, pause

Abstractive Summarization (of Customer Reviews)

Analysis of Evaluative Documents

- Conventional multi-document summarization focuses on factual info, e.g., news
- Rapid accumulation of evaluative documents on the web
  - Expressing the author's subjective sentiments, e.g., like or not like, good or bad
- Sentiment analysis focuses on extracting these opinions (e.g., Bing Liu's tutorial)
- Our work goes beyond extraction to produce a summary

Our Multimedia Approach



Sample review about a camera

…… the canon computer software used to download , sort , . . . is very easy to use. the only two minor issues i have with the camera are the lens cap (it is not very snug and can come off too easily). . . .

the menus are easy to navigate and the buttons are easy to use. it is a fantastic camera . . . .

Extracting Evaluative Features

A. Which **features** of the entity are evaluated in the reviews?

B. What is the **polarity** of each feature? (positive or negative)

C. What is the **strength** of each feature? (rather good vs. extremely good, [-3 .. +3]

[Hu, Liu 2004; Wilson et al. 2004, Etzioni 2005]

Example: Extracting Features

…… the canon computer software used to download , sort , . . . is very easy to use (+2). the only two minor issues i have with the camera are the lens cap (it is not very snug (-2) and can come off too easily (-2)). . . .

the menus are easy to navigate(+1) and the buttons are easy to use(+1). it is a fantastic(+3) camera …

Extractive vs Abstractive Summaries

- Extractive: extracts a representative subset of the sentences from the original corpus
  - Fluent and varied language (c.f. DUC 2005)
  - Intuitively not well suited to summarize evaluative corpus
  - Can't express *distribution* of opinions (‘some/all’)
  - Can't *aggregate* opinions either numerically or conceptually
- Abstractive: generates text that expresses the most relevant info in the original corpus
  - Possibly not fluent and repetitive
  - Intuitively well suited to summarize evaluative corpus

Generating an Abstractive Summary

- Map extracted features onto a taxonomy of product features at different levels of abstraction
- Such a mapping:
  - **Eliminates redundancy**
  - **Provides a conceptual organization of the features**
  - **Increases user familiarity with the features**

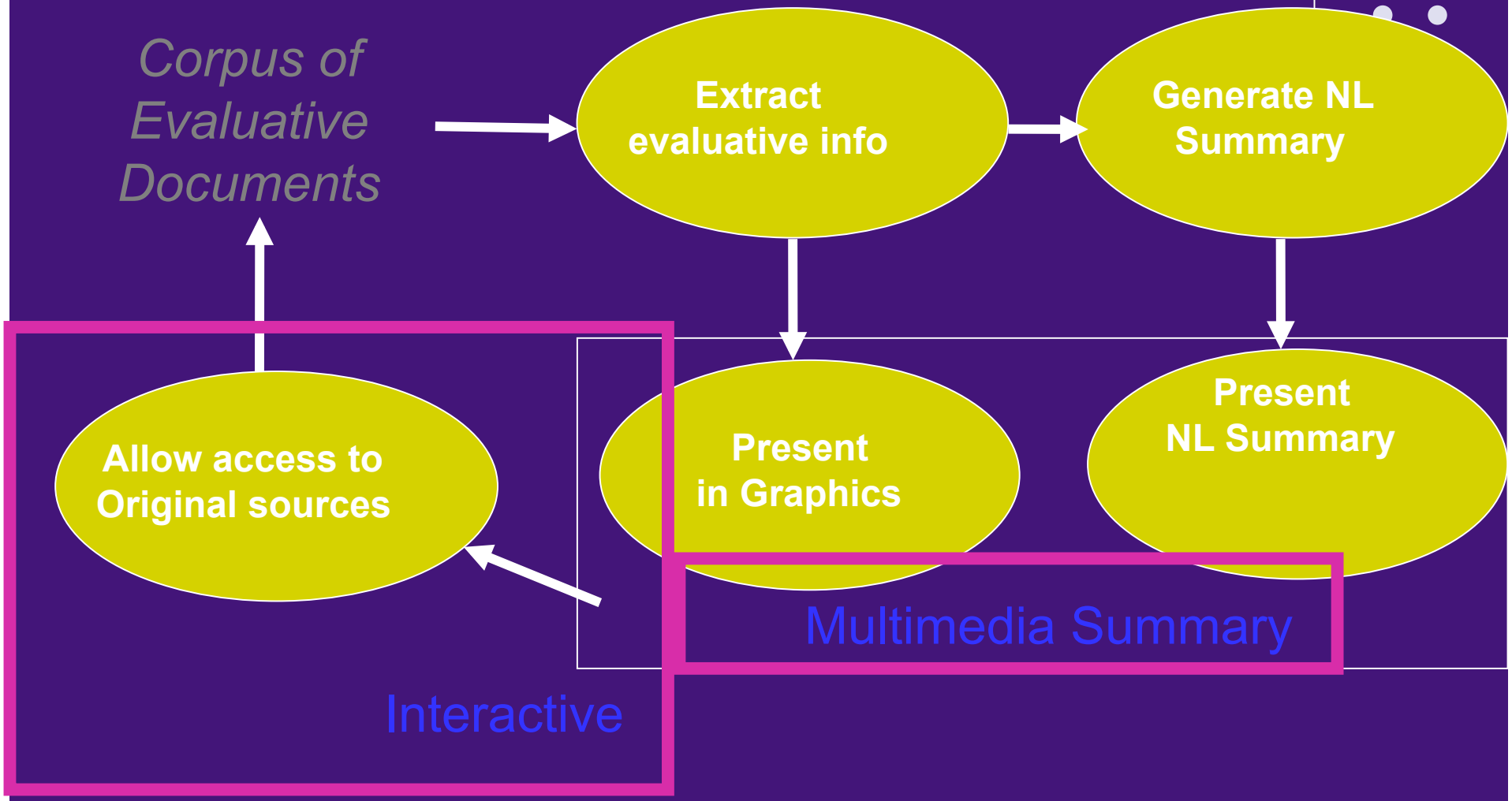
Extracted Features after Mapping

Canon G3 PS Digital Camera **[-1,-1,+1,+2,+2,+3,+3,+3]**

1. User Interface **[+2]**
   - Button **[+1]**
   - Menus **[+2,+2,+2,+3+3]**
   - Lever **[ ]**
2. Convenience **[ ]**
   - Battery **[ ]**
     - Battery life **[-1,-1,-2]**
     - Battery charging system **[ ]**
   - ….
3. ….

Each list gives the polarity and strength of the feature

Our Multimedia Approach



Ongoing and Future Work

Reminder: Key Research Questions

2. How to summarize the text:
   - From a single modality, e.g., meeting notes, emails, customer reviews, blogs?
   - From across multiple modalities/genres? Is there a universal set of features, or are the features genre-specific?

Multi-modal: Meetings & Emails

- We focused on meetings and emails and aim to develop summarization techniques that are effective in both domains
- We identified a set of features: cross-fertilization turns out to be valuable, i.e., these are features effective for both modality
  - Domain adaptation proved effective and can be a general technique in the absence of annotated data

Abstracting Conversations

1. Interpretation
   - Used a conversation ontology written in OWL and sentence-level classifiers, identifying phenomena such as action items, decisions, problems
2. Transformation
   - Sentences are weighted according to their ontology links
   - ILP finds the subset of sentences that maximizes a scoring function, given a length constraint
3. Generation of the text abstracting the chosen sentences

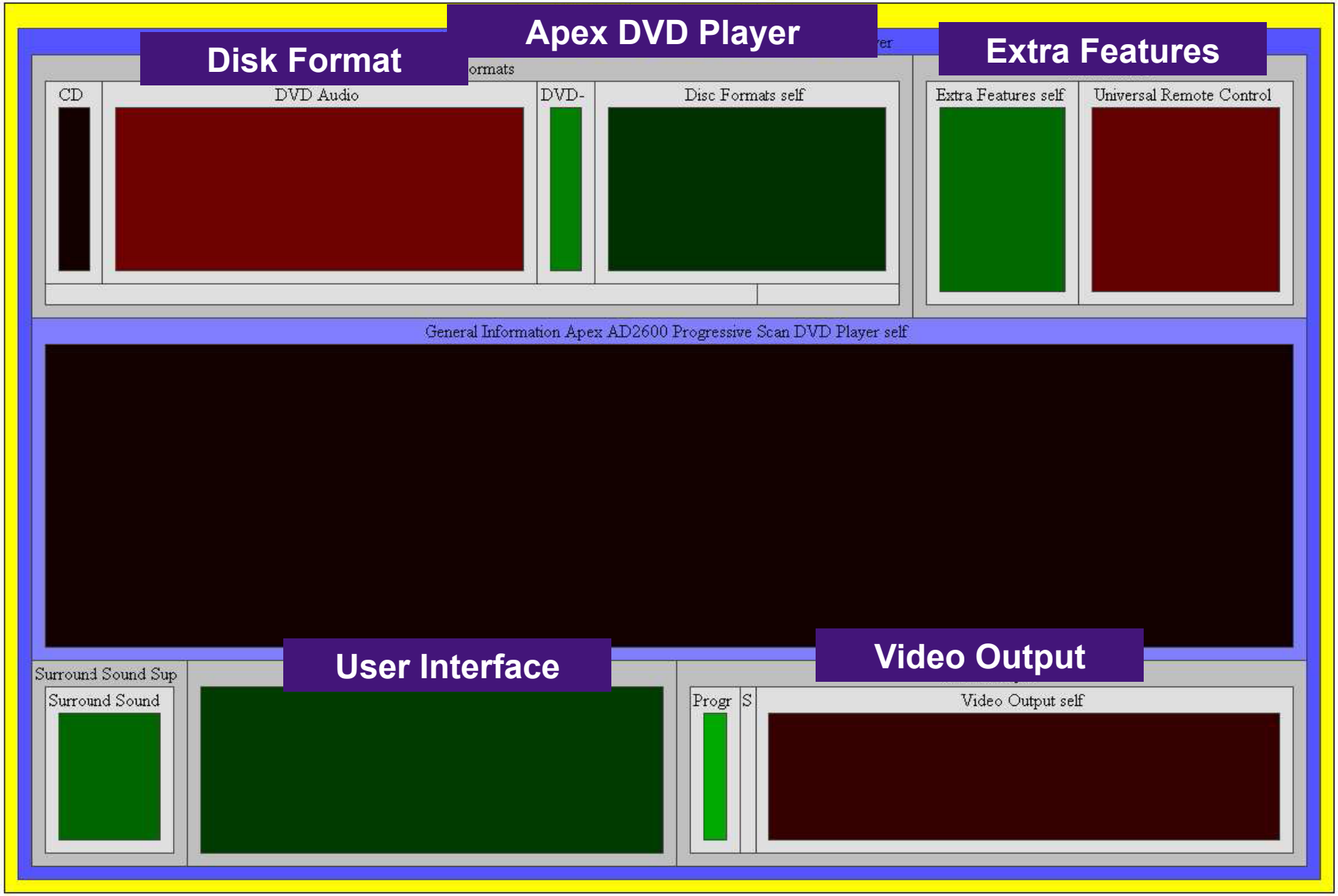
Summarizing Blog Opinions

- From our perspective, this task combines a lot of our interests and what we have done
  - Conversation
  - Subjectivity & Polarity
  - Summarization
- An abstractive summarizer of blog opinions is being built

Future Work

- Deeper analysis on subjectivity, e.g., contextual cues, detecting out-of-topic sentences
- Identification of topics (and when they change) in a conversation
- Automated recognition of a conversation across multi-modal documents still a big open problem
- Visualization tools for conversations

Thank You!

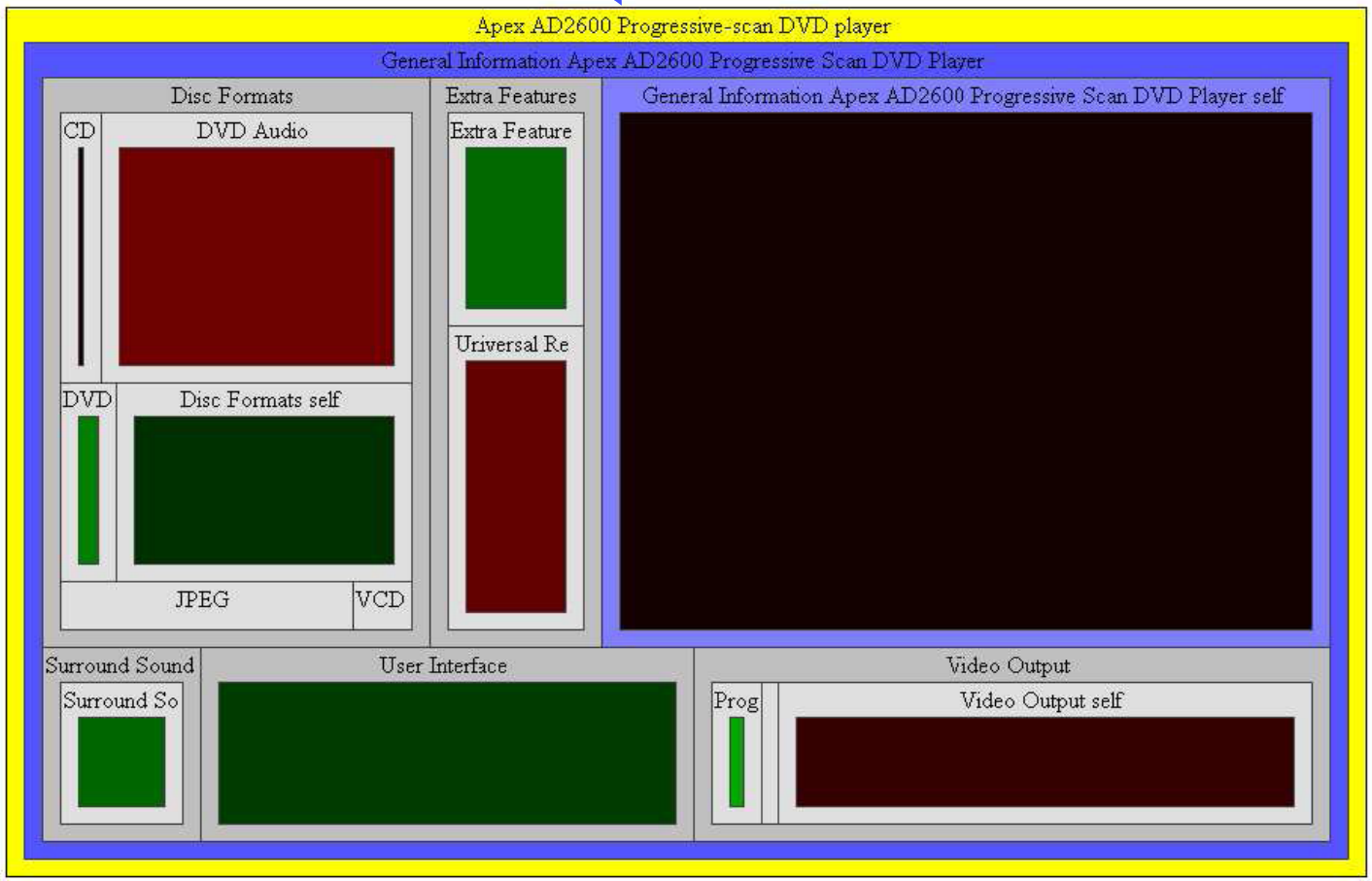

Apex DVD Player
Disk Format
Extra Features
ormats
CD
DVD Audio
DVD-
Disc Formats self
Extra Features self
Universal Remote Control
General Information Apex AD2600 Progressive Scan DVD Player self
User Interface
Video Output
Surround Sound Sup
Surround Sound
Progr
S
Video Output self

Textual Summary

Graphical Summary

Summary of customer reviews for: Apex AD2600 Progressive-scan DVD player

Most customers disliked the Apex AD2600 [1]. Although many customers found the user interface [2] to be good, many users thought the available video outputs [3] was poor. However, many users liked the range of compatible disc formats [4], even though many customers found the compability with DVD audio [5] discs to be very poor.



For the price , it 's a very nice dvd player . The front door is miss aligned on my unit and you have to manually life it up just so slightly for the door to close , a very annoying thing after ahwile . It does play a wide range of formats as advertised which is very nice . And so far have not had any problems with dvds not being able to play . Recommended to anyone looking to purchase a low priced dvd player and not expecting any bells or whistles from a brand name one like sony .

Original Review(s)

An Example of Our Abstractive Summary

"Almost all users loved the Canon G3 possibly because some users thought the physical appearance was very good. Furthermore, several users found the manual features and the special features to be very good. Also, some users liked the convenience because some users thought the battery was excellent. Finally, some users found the editing/viewing interface to be good despite the fact that several customers really disliked the viewfinder."

Example Taxonomy

1. Canon G3 PS Digital Camera [*canon,canon PS g3, digital camera, camera*,…]

1.1 User Interface [*button, menus, lever*]

1.2 Convenience [ ]

- Battery [*battery life, battery charging system, battery*]
- Self Timer [ ]
- Burst Mode [*speed, mode*]
- Rapid Fire Shot [*speed*]
- Delay Between Shots [*unresponsiveness, delay, speed, lag time, lag*]
- ….

2. Not Placed *[manual, function, quality, strap, service, shoot, learning curve,…]*

Reminder: Business Intelligence Scenario

- In a meeting, the VP on marketing raised the topic of developing a new product
- Subsequent emails continued the discussion on:
  - Web documents describing similar products;
  - User reviews on those products.
- This conversation spans meeting notes, emails, web documents, customer reviews and blogs
- How to automate the generation of a concise summary of the conversation?

Discussion: What can We Do together?

- Pilot data sets?
  - Emails, e.g., bugs reports, customer requests? Blogs?
  - Even better if there are already generated summaries
- Tasks for which summarization of text is valuable?
  - Extractive and/or abstractive
- Provide feedback to test or train our techniques?
- Share or develop ontologies?

Interpretation

- We create an ontology of conversations
  - Written in the Web Ontology Language (OWL)
  - Representing participants, entities, dialogue acts, etc.
- We use multiple sentence-level classifiers, identifying phenomena such as *action items, decisions*, and *problems*.
- These classifier outputs are used to populate the ontology with low-level & high-level instance data
  - Low-level – e.g. Individual sentences and their properties
  - High-level – e.g. Messages based on repeated patterns

Transformation

- We select the most informative content for which we will generate novel abstractive text
- This is done via an optimization framework, using Integer Linear Programming
- Sentences are weighted according to their ontology links and the entities they contain (e.g. “interface,” “budget”, “LCD screen”)
- ILP finds the subset of sentences that maximizes a scoring function, given a length constraint

Intermediate Results

- This optimization framework outperforms several greedy sentence selection approaches
- The selected sentences are highly informative to generate abstract summaries
- The next step, and our current focus, is on grouping (and abstracting over) selected sentences and entities to generate novel text