Generating Summaries of Line Graphs

Priscilla Moraes, Gabriel Sina, Kathleen McCoy and Sandra Carberry
Department of Computer and Information Sciences
University of Delaware, Newark, Delaware, USA
[pmoraes | gsina | mccoy | carberry]@udel.edu

Abstract

This demo presents a Natural Language Generation (NLG) system that generates summaries of informational graphics, specifically simple line graphs, present in popular media. The system is intended to capture the high-level knowledge conveyed by the graphic and its outstanding visual features. It comprises a content selection phase that extracts the most important content of the graphic, an organization phase, which orders the propositions in a coherent manner, and a realization phase that uses the text surrounding the article to make decisions on the choice of lexical items and amount of aggregation applied to the propositions to generate the summary of the graphic.

1   Introduction

Multimodal documents from online popular media often contain information graphics that augment the information found in the text. These graphics, however, are inaccessible for visually impaired users or in environments where the image cannot be processed/displayed. Our system captures the high-level content of the graphic and produces a textual summary that conveys it. Figure 1 shows the system architecture.

The first step is the identification of the presence of a graphical image in the web page by a Browser Helper Object (BHO) (Elzer et al., 2007). If a graphic is present on the web page, the Graphical Information Extraction Module (VEM) (Chester & Elzer, 2005) is triggered by the BHO in order to extract the data from the image. The VEM then produces an XML representation of the graphic that is used by the Intention Recognition Module (IRM) for simple bar charts (Elzer, Green, Carberry, & Hoffman, 2006), simple line graphs (Wu, Carberry, Elzer, & Chester, 2010) and grouped bar charts (R. Burns, Carberry, & Elzer, 2010; R. Burns, Carberry, & Schwartz, 2013; R. J. Burns, 2013). The XML representation

of the graphic, along with the intended message identified by the IRM, is sent to the Generation Module (GM), which produces a textual summary of the most important content presented in the graphic. The system produces an initial summary and follow-up responses for simple bar charts (Demir, Carberry, & Elzer, 2009; Demir, Carberry, & McCoy, 2008) and this demo presents the GM for simple line graphs.

This demo focuses on presenting the generation phase of the system. For that, we will demonstrate the generation of summaries in the context of a digital library that is available online[1] and that contains information graphics collected from online popular media, along with the articles containing the graphics. In addition, we have included hand-generated XML representations for the graphics (the current VEM is not fully robust). For each article that contains a graph, the user can choose to have access to the generated summary by clicking on the "Generate summary" button (highlighted in Figure 2). Figure 2 shows a screenshot on which the graph shown in Figure 3 has its article featured.

For accessibility projects that may use our system (applications developed for visually impaired users, for example), the application might use a combination of key strokes to allow user interaction. The module of the system that is the focus of this demo is the Generation Module.
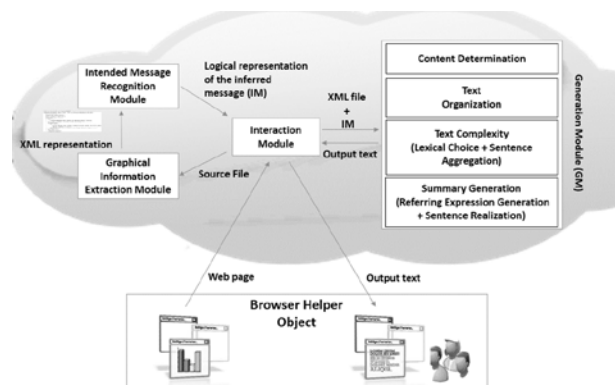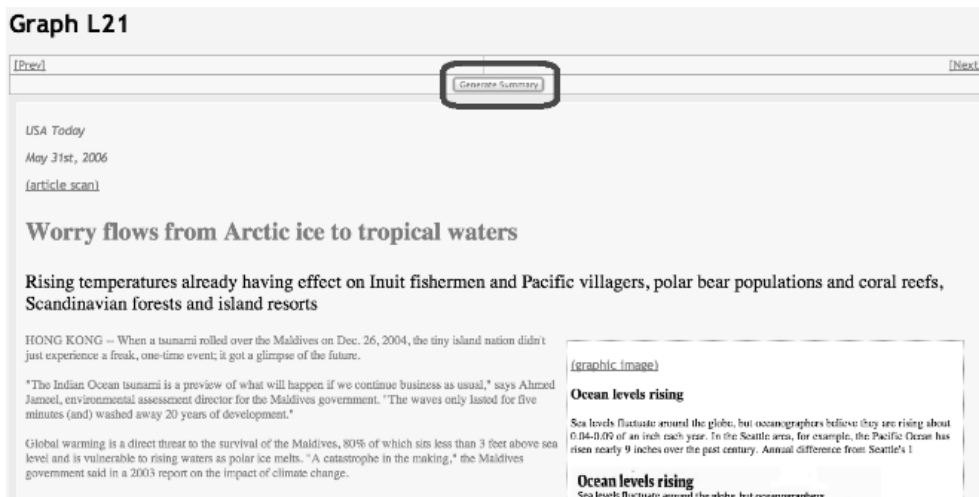

Figure 1: System Architecture

---

[1] http://ir.cis.udel.edu/~moraes/udgraphs

Figure 2: Digital library screenshot where we have added summary generation functionality.

2 Generation Module

For generating summaries of line graphs, the first step is the selection of content. In order to select the most important features of the line graph that should be conveyed in the summary, the system represents the intended message and the visual features identified by a human subject experiment (Greenbacker, Carberry, & McCoy, 2011) using a graph. A centrality-based algorithm, which is an adapted version of PageRank (Page, Brin, Motwani, & Winograd, 1999), is then implemented to select the most important information (represented as nodes in the graph). This implementation allows semantic relationships between propositions to be represented on the edges of the graph. The core of the content selection framework is to detect present outstanding visual features in the graphic, along with its intended message, in order to select nodes. Details in the content selection phase are available in the work presented at (P. S. Moraes, Carberry, & McCoy, 2013).

The next phase is the organization of the selected content. The organization phase works by ordering the selected propositions such that the delivered summary is fluent and coherent. The summaries are organized having an introduction section, a detailed section and a conclusion. The introduction consists of overall information about the line graph (the type of the graph, the entity being measured, the volatility of the graph and its intended message). The identified trends are described in the detail section. For this part of the summary, pieces of the graphic that outstand due to its visual features may be described first, being followed by other trends. Finally, the conclusion section of the summary presents computational information about the graphic (overall value and rate change, time span of the graphic, maximum and minimum points and dates when they occur). The strategies on organizing the summaries are described in (P. Moraes, McCoy, & Carberry, 2014).

The last step of the Generation Module is the aggregation of propositions into more complex sentences. This decision is usually left to the designer's choice on how much aggregation to perform when generating text. Some systems are designed to generate simple text for people with low reading abilities (Williams & Reiter, 2005a). As stated by (Williams & Reiter, 2005b), most NLG systems available generate text for high-skilled users. Our system generates line graph summaries that fit the reading level of the article in which the line graph appears. We contend that users generally read articles from venues they feel comfortable with reading. In this manner, we intrinsically assess the user's reading level without needing to actively survey it.
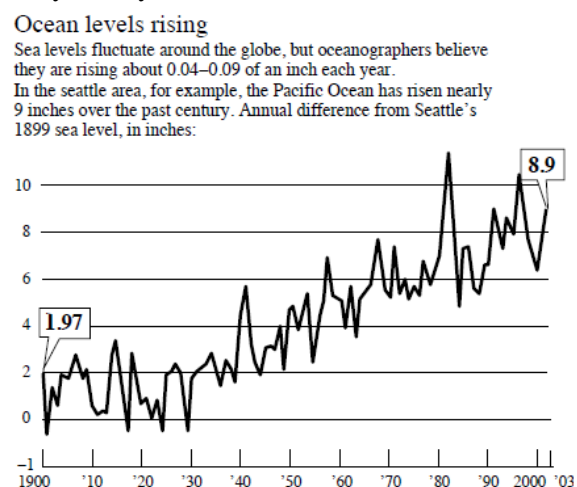

Figure 3: A line graph present in popular media.

The first step of the aggregation phase is to assess the reading level of the article's text. There is a myriad of techniques to measure the reading level of text. Much of them use machine learning techniques in order to learn text constructions and lexicalization used in different grade levels. As presented in (P. Moraes et al., 2014), simpler and well established reading level measurement techniques suffice for our scenario. The work shows that Flesh-Kincaid (Kincaid, Fishburne, Rogers, & Chissom, 1975) and SMOG (Laughlin, 1969) provide the set of information needed by the system in order to make decisions of syntactical text complexity.

After assessing the reading level of the article, the system then uses the text plan that applies to the identified reading level. Text plans define rules on Noun Phrase (NP) density and lexical choice. When describing an entity, attributes of this entity can be added to the NP as modifiers using either adjectives e.g. "*a highly volatile rising trend*", conjunctions e.g., "*the rising trend is volatile and steep*" or relative clauses e.g. "*a rising trend, which is highly volatile*". When the modifier of an NP is a Verb Phrase (VP), it is combined using a relative clause e.g., "*the line graph, which presents the number of jackets sold in 2013...*" VPs can be modified by adverbs e.g., "*the falling trend is **very** steep*". The text plans apply rules within sets of propositions that are grouped hierarchically. The system then uses the appropriate lexical items (*highly volatile vs ups and downs; conveys vs shows*) and applies the appropriate amount of aggregation in order to realize sentences.
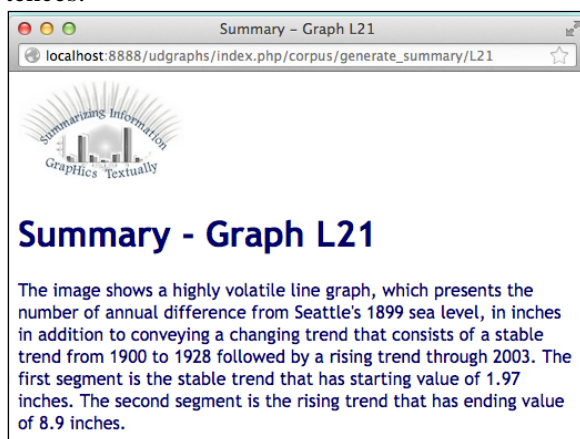

Figure 4: Pop up window with the resulting summary generated by the system.

Figure 4 and Figure 5 display the summaries generated for a user whose reading level is 11th-13th grade and 5th-7th grade respectively. From these one can see the different aggregation and lexical choice decisions made for the different reading levels. The system also includes appropriate pronominalization in order to avoid repetition of the referring expressions (P. Moraes et al., 2014).

```
Summary for Grades > 5 and <= 7
The image shows a line graph.
The line graph has ups and
downs. It presents the number
of annual difference from
Seattle's 1899 sea level, in
inches. It conveys a changing
trend. It consists of a stable
trend from 1900 to 1928
followed by a rising trend
through 2003. The first segment
is the stable trend. It has a
starting value of 1.97 inches.
The second segment is the
rising trend. It has an ending
value of 8.9 inches.
(SMOG 4.8)
```

Figure 5: Example of a summary adapted to the reading level of grades 5 to 7.

For the surface realization phase we use FUF/SURGE (Elhadad & Robin, 1999) to create the templates for realization. The template are created based on the text plans defined for a given reading level, as described above.

3 Conclusion

This paper presents the demonstration of the generation module of SIGHT. For the demo, the generation module works on a digital library that archives informational graphics collected from popular media available online. The aggregation phase of the generation module tailors the syntactical complexity of the generated text to that of the article's text in which the graphic appears.

An evaluation of the text summaries generated at different reading level is presented at (P. Moraes et al., 2014). It shows that, indeed, different users have different preferences regarding different text designs.

4 Future Work

A more automated way of defining a text plan for a given reading level is under investigation. We will explore techniques for learning how different text constructions can affect reading measures and then using these learned models when choosing an

adjective over a relative clause for increasing the NP density and use of passive voice, for example.

Choosing lexical items that are classified by age is another possibility. We plan on investigating how the usage of word frequency by age/grade level (Carroll, 1972) might influence the overall generated summaries.

5 Acknowledgement

References

Burns, R., Carberry, S., & Elzer, S. (2010). *Visual and spatial factors in a bayesian reasoning framework for the recognition of intended messages in grouped bar charts.* Paper presented at the Proceedings of the AAAI Workshop on Visual Representations and Reasoning.

Burns, R., Carberry, S., & Schwartz, S. E. (2013). *Modeling a Graph Viewer's Effort in Recognizing Messages Conveyed by Grouped Bar Charts.* Paper presented at the UMAP.

Burns, R. J. (2013). *Automated intention recognition of grouped bar charts in multimodal documents.* University of Delaware, Ann Arbor. Retrieved from http://search.proquest.com/docview/1318643227?accountid=10457

Carroll, J. B. (1972). A New Word Frequency Book. *Elementary English, 49*(7), pp. 1070-1074.

Chester, D., & Elzer, S. (2005). *Getting computers to see information graphics so users do not have to.* Paper presented at the the Proceedings of the 15th International Symposium on Methodologies for Intelligent Systems.

Demir, S., Carberry, S., & Elzer, S. (2009). Issues in Realizing the Overall Message of a Bar Chart. In N. Nicolov, G. Angelova & R. Mitkov (Eds.), *Recent Advances in Natural Language Processing V* (pp. 311-320): John Benjamins.

Demir, S., Carberry, S., & McCoy, K. F. (2008). *Generating textual summaries of bar charts.* Paper presented at the Proceedings of the Fifth International Natural Language Generation Conference, Stroudsburg, PA, USA.

Elhadad, M., & Robin, J. (1999). SURGE: a comprehensive plug-in syntactic realization component for text generation. *Computational Linguistics*.

Elzer, S., Green, N., Carberry, S., & Hoffman, J. (2006). A Model of Perceptual Task Effort for Bar Charts and its Role in Recognizing Intention. *International Journal on User Modeling and User-Adapted Interaction, 16*(1), 1-30.

Elzer, S., Schwartz, E., Carberry, S., Chester, D., Demir, S., & Wu, P. (2007). *A Browser Extension For Providing Visually Impaired Users Access To The Content Of Bar Charts On The Web.* Paper presented at the the Proceedings of the International Conference on Web Information Systems and Technologies.

Greenbacker, C., Carberry, S., & McCoy, K. (2011, July). *A Corpus of Human-written Summaries of Line Graphs.* Paper presented at the Proceedings of the UCNLG+Eval: Language Generation and Evaluation Workshop, Edinburgh, Scotland.

Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel.

Laughlin, G. H. M. (1969). SMOG Grading-a New Readability Formula. *Journal of Reading, 12*(8), pp. 639-646.

Moraes, P., McCoy, K., & Carberry, S. (2014). *Adapting Graph Summaries to the Users' Reading Levels.* Paper presented at the Proceedings of the 8th International Natural Language Generation Conference.

Moraes, P. S., Carberry, S., & McCoy, K. (2013). *Providing access to the high-level content of line graphs from online popular media.* Paper presented at the Proceedings of the 10th International Cross-Disciplinary Conference on Web Accessibility, Rio de Janeiro, Brazil.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank Citation Ranking: Bringing Order to the Web: Stanford InfoLab.

Williams, S., & Reiter, E. (2005a). *Appropriate Microplanning Choices for Low-Skilled Readers.* Paper presented at the IJCAI.

Williams, S., & Reiter, E. (2005b). *Generating readable texts for readers with low basic skills.* Paper presented at the Proceedings of the 10th European Workshop on Natural Language Generation (EWNLG 2005).

Wu, P., Carberry, S., Elzer, S., & Chester, D. (2010). *Recognizing the intended message of line graphs.* Paper presented at the Proceedings of the 6th international conference on Diagrammatic representation and inference, Berlin, Heidelberg.