Semantic Modeling of Multimodal Documents for Abstractive Summarization

Charles F. Greenbacker[1], Kathleen F. McCoy[1], Sandra Carberry[1], and
David D. McDonald[2]

[1] Department of Computer and Information Sciences, University of Delaware,
Newark, Delaware, USA* {charlieg,mccoy,carberry}@cis.udel.edu
[2] SIFT, LLC, Boston, Massachusetts, USA dmcdonald@sift.info

Abstract. We describe a method for semantic modeling of multimodal
documents and discuss how this can be used to generate an abstractive
summary. Information extracted from the text by a semantic parser and
from the graphics by a graph understanding system is combined into a
single knowledge base. By operating at the semantic (rather than the
surface) level, we are able to integrate information collected from both
text and non-text sources. From this unified semantic model, we can
evaluate the importance of each part of the extracted knowledge and
produce a comprehensive summary of the entire multimodal document.

1 Introduction

This work is part of a larger ongoing effort to produce better and more inclusive descriptions of the information contained in multimodal documents found
in popular media. Multimodal documents consisting of text and information
graphics (such as bar charts and line graphs) pose a difficult challenge for traditional natural language processing techniques. The graphical content is not
always duplicated in the text of the document [4], and yet the graphic may
contain valuable information important to the article's message. The content
creator had a reason for including the graphic in the multimodal document, and
if the graphic is ignored, the summary may not be a good representation of the
document as a whole. Our current focus is on combining information extracted
from the text with the most important information conveyed by the graphics in
order to produce an integrated summary of the entire article. This line of work
helps address what is commonly known as the *information overload* problem by
condensing the information contained in multimodal documents into brief synopses. This is particularly important for people with visual impairments, due
to the significant time investment required for them to read lengthy articles, as
well as the additional difficulties they face in accessing graphical content. We approach summarization from a generation perspective, thus our goal is to produce
a natural language summary as output.

2 Motivation

A summary which considers the information contained in graphical sources should be abstractive in nature. Most summarization tools utilize extractive techniques [19, 20], whereby the most important sentences are extracted from a document and then reassembled to form the summary. However, this approach cannot faithfully retrieve the information stored in graphics since these non-textual modalities offer no sentences for extraction. Some research into summarizing or otherwise representing the content of a graphic has relied solely on captions and other sentences in the article explicitly referring to the graphic in order to summarize it [2, 34]. However, studies have shown that the graphical content is often not repeated in the accompanying article text [4] and captions are often uninformative [14]. Work on summarizing multimodal documents has taken images and text into account to some extent, by doing very shallow processing on an image to categorize it [11], or using the accompanying text to disambiguate image contents [31], but none that we are aware of consider a graphic on par with text in terms of adding communicative content to a document. Furthermore, summaries produced by extractive methods in general, while syntactically correct, have been shown to lack cohesion and suffer from ambiguity and referent identification issues [26]. In contrast, an abstractive summary would address both of these issues by working from an underlying semantic representation of the text and graphics, and by using natural language generation techniques for text structure and surface realization to ensure text coherence.

One possible approach to facilitate extractive summarization of multimodal documents would be to first generate a textual description of the graphics [12, 7] and then insert this description into the document text before performing sentence extraction. However, not only would the resulting summary suffer from the limitations inherent to extractive methods described above, it would face additional difficulties because the combined text (machine-generated graphical description inserted into original text of article) would be written by two different authors in two different styles, thus leading to even more coherence issues. Therefore, not only do graphics require an abstractive treatment, information from both text and graphics should be semantically integrated in order to generate a cohesive summary of the entire multimodal document.

Automatic summarization methods that more closely approximate the human process of conceptual integration and regeneration in writing an abstract will likely produce results which are more human-like than that of traditional extraction techniques. However, the automatic abstractive summarization of text has proven to be quite a challenging problem [27], even without considering the incorporation of multimodal sources of information. Efforts directed towards abstraction have included the modification (i.e., editing & rewriting) of extracted sentences [18], as well as using partial semantic analysis with text regeneration and elaboration to produce indicative-informative abstracts from technical information [30]. Some research into "semantic abstraction summarization" has aimed to represent the summarized content as a graphical condensate [17], rather than producing a natural language summary. Our work shares similarities with

the knowledge-based text condensation model of Reimer & Hahn [29], as well as with Rau et al. [28], who developed an information extraction approach for conceptual information summarization, though our goal is to represent both the text and the graphics in a single conceptual model in order to generate a natural language summary of a multimodal document.

3 Methodology

In the remainder of this paper, we will present our method for extracting information from text and graphical sources to build a semantic model that captures the information content of both the text and the graphics, and then discuss how an abstractive summary can be produced from this model.

3.1 Text Understanding

The semantic parsing of document text is performed by Sparser [21], a bottom-up, phrase-structure-based chart parser, optimized for semantic grammars and partial parsing.[3] While most parsers stop at a structural description, Sparser produces a disambiguated conceptual model. It outputs categorized objects and relationships, creating and adding specific facts to instances of highly-structured, predefined prototypes. Sparser contains a built-in, sophisticated linguistic model of core English grammar, as well as a model of common items such as names, locations, times, and amounts. Given a document and domain-specific grammar, Sparser performs a linguistic analysis, identifying each part of the text where the subjects of its grammar appear, and emitting partially-saturated referents (PSRs) as a semantic representation of what it recognizes [23]. A PSR is a semantically-incomplete representation of a concept for which some of the characteristic information can be missing; in other words, an object possibly lacking values for some of its attributes.

Existing Sparser grammars provide coverage for several different domains, including business news articles. A collection of multimodal documents from popular media has been assembled, most of which contain article text accompanied by information graphics. Among these articles are many in the business news domain. We have extended Sparser's semantic grammar for this domain, allowing it to analyze texts like the article entitled "Will Medtronic's Pulse Quicken?" from the May 29, 2006 edition of Businessweek magazine.[4] Such texts convey information about stock prices, earnings forecasts, analysts' predictions, and market conditions. Sparser recognizes these semantic entities, builds and modifies PSRs to represent them, and adds these to the semantic model being constructed.

3.2 Graph Understanding

As image understanding research has not yet developed tools capable of extracting semantic content from every possible image, we must restrict our focus to a

---

[3] https://github.com/charlieg/Sparser
[4] http://www.businessweek.com/magazine/content/06_22/b3986120.htm

limited class of images for the prototype system implementation. We have opted to leverage capabilities developed for the SIGHT system [9], which generates textual summaries of information graphics found in popular media (e.g., magazines, newspapers) for people who have visual impairments. Rather than focusing on specific data points or the shape of the graphic (as might be appropriate for a scientific graph), SIGHT conveys the underlying message (made apparent by the choice of graph type and the communicative signals entered into the graphic by the graph's author) along with propositions that are highlighted by visual features. For example, given the bar chart in Fig. 1, SIGHT generates the following initial summary [8] in about one minute on a modern PC:

> *Following a moderate rise between the year 1993 and the year 1994, the graphic shows a decreasing trend in the amount of newark rainfall for july over the period from the year 1994 to the year 2002. The amount of newark rainfall for july shows the largest drop of about 1.29 inches between the year 1999 and the year 2000. With the exception of a few rises, slight decreases are observed almost every year over the period from the year 1994 to the year 2002.*

Our framework is general enough to accomodate arbitrary image types and other modalities (e.g., audio, video), however. Incorporating other modalities would require adding a module capable of mapping the particular modality to its underlying message-level semantic content.
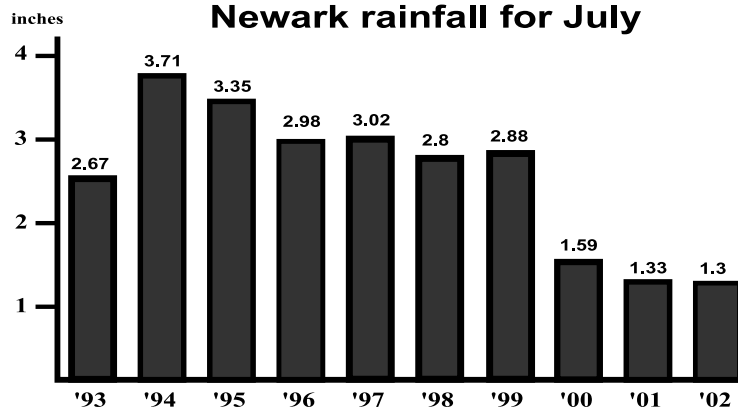

Fig. 1. Example bar chart processed by SIGHT.

Several modules of the SIGHT system are relevant to our current work. The image file is first analyzed by SIGHT's visual extraction module [6], which produces an XML representation of the information stored in the graphic [16]. For example, given a bar chart, the XML output contains axis labels, information about each bar (e.g., position, height, value, color/shading), captions and legends, etc. We contend that this "raw information" extracted from the graphic

(the *visual* level) is not the proper level of understanding upon which to base a summary of the article. Far more pertinent is the communicative intent of the graphic as it relates to the overall document (the *message* level). Thus, SIGHT's intention recognition module [15] applies an inference model (including Bayesian networks) to reason about the communicative signals contained in the graphic (based on attributes derived from statistical tests, cognitive psychology research into perceptual task effort, and visual features) to identify the intended message of a bar chart (e.g., rising trend, rank of an entity). Recent work has extended this to line graphs [32, 33] and a subclass of grouped bar charts [3].

Once the communicative intent has been identified, the system extracts additional salient propositions that expand on the graph's intended message. On the basis of human subjects experiments, the propositions are marked with varying levels of importance depending on the intended message and visual features of the graph. These propositions, along with the intended message, represent the knowledge conveyed by the graphic and capture the *message-level* content that the graph should contribute to the summary of the document. The propositions capture a variety of concepts, including time span, degree of volatility, exceptions in trends, and entity comparisons. The inferred message and extracted propositions are added to the semantic model, making connections to concepts previously derived from the text as appropriate. The SIGHT system is already capable of extracting the most salient propositions from simple bar charts [7], and current efforts are working to extend this capability to line graphs and grouped bar charts as well.

3.3 Knowledge Representation

For our knowledge representation system, we use KRISP ("Knowledge Representation In SParser"): a system of typed, structured objects organized under a foundational ontology [22]. The PSRs recognized by Sparser are stored in KRISP as instantiations of pre-defined categories in a model. As Sparser obtains more and more information about a particular object, the corresponding entry in the model becomes more complete (i.e., "filled-out" or "saturated"). In addition, meta-information relating to the concept, such as document structure (e.g., position in the text) and the use of rhetorical devices (e.g., appearance in a comparison by means of juxtaposition), is included in the model as well. Finally, the model stores the original phrasings used in the source document to express each concept in the form of tree-adjoining grammar (TAG) derivation trees, which are the underlying syntactic representation for Sparser; these phrasings are made available for use during the generation phase.

The semantic model of the text built by Sparser is extended to cover the entire multimodal document by decomposing the intended message and propositions extracted from the graphics by SIGHT and inserting this information into the model. Though the graphs often contain material not repeated in the text, there is usually a high degree of connectedness between concepts presented in the text and those in the information graphics. This is represented in the model by instantiating the new objects and relationships introduced by the graphs,

forging new connections to existing entries, and filling the slots of previously-observed PSRs as appropriate. In addition, mirroring the document structure and rhetorical device details associated with the text-based concepts, the propositions extracted from the graphic are marked with importance values derived from the human subjects experiments. These ratings are influenced by the intended message and visual characteristics of the graph.

Sample Semantic Model Figure 2 offers a high-level overview of the semantic model built for the Medtronic article mentioned in Sect. 3.1, while Fig. 3 provides a detailed look at a zoomed-in section of the same model. Each node in Fig. 2 represents an individual concept recognized in the document either by Sparser or the graph understanding component. The name indicates the conceptual category with a number to distinguish between instances. In the interest of space, individual attributes of model entries have been omitted from this diagram, but are available in Fig. 3. Lines connecting nodes indicate a semantic link between the corresponding concepts (i.e., one fills an attribute slot of the other). In addition to entities from the text recognized by Sparser, this diagram also shows the overall intended message (ChangeTrend1) and informational propositions (e.g., Volatile1) the SIGHT analyzer extracted from the graphic. This way, information gathered from text and graphical sources can be integrated at the *semantic* level regardless of the format of the source.
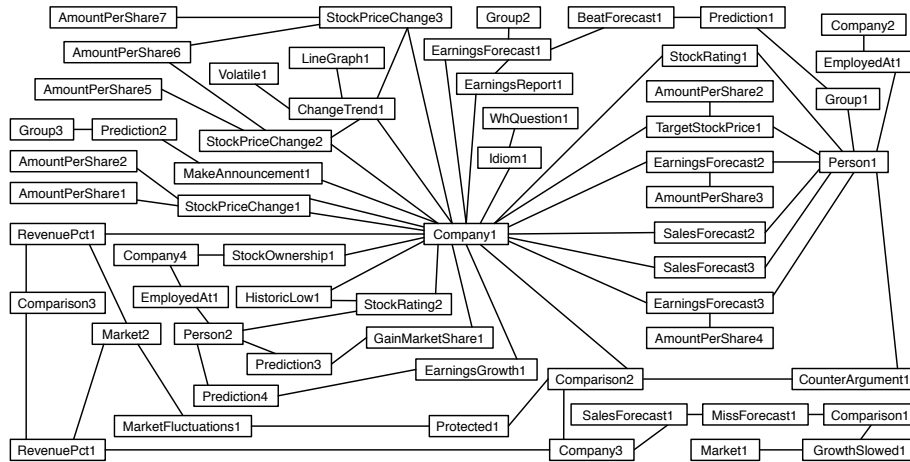
Fig. 2. High-level (low-detail) representation of semantic model for Medtronic article.

Figure 3 zooms into a portion of the model to show more detail for individual concepts. The top section of each box contains the category name and instance number, the middle section shows various attributes with their values (if any), and the bottom section lists the original phrasings expressing these concepts
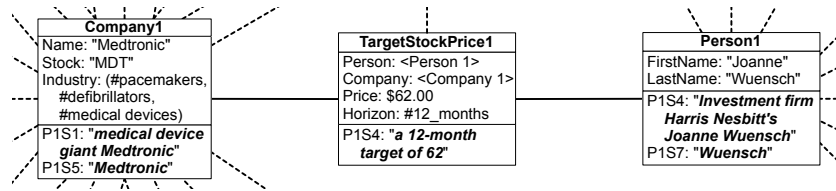
Fig. 3. Low-level detail of model showing concepts used in Sect. 4 sample output.

(formally stored as a synchronous TAG). Attribute values in angle brackets ($<>$) are references to other concepts. Hash marks (#) denote a symbol that has not been fully instantiated as a concept in the current model.

Encapsulating Document Structure The model also tracks various details regarding document structure. Each recorded expression is marked with a sentence tag (e.g., "P1S4" stands for "paragraph 1, sentence 4" as seen in Fig. 3), indicating exactly where each concept appeared in the text. This allows the content rating metrics to take into account the location of a referent, whether mentioned in the title or buried in the middle of a paragraph, when determining salience. Information obtained from graphical sources receives a similar treatment: entries are marked with importance values derived from our analysis of the corresponding propositions (e.g., due to their rating in our human subjects experiments). As such, the fact that a particular concept is featured prominently in an information graphic is considered during content rating. Certain rhetorical devices that highlight a concept's significance are accounted for as well and represented as distinct entries in the semantic model (e.g., Comparison2 and Idiom1 in Fig. 2). We can accomodate documents of any length, limited only by the storage and processing capacities of the computing environment. Dealing with longer documents is not necessarily more difficult than shorter ones. Articles with a high degree of focus on a central theme tend to result in elaborating and extending existing concepts, rather than introducing new ones. As a result, the corresponding semantic model can increase in detail ("saturation level") more so than in size. Additionally, the model can be adapted to accomodate information from multiple documents by inserting and connecting new concepts while tracking their source, thus facilitating multi-document summarization.

Enhancing Expressibility Although they are represented in Fig. 3 as strings, the original expressions used to realize the PSRs recognized by Sparser are stored in the semantic model as parameterized synchronous TAG derivation trees. These trees are used as the "raw material" for realizing the corresponding referents and relationships in text during the generation phase [24]. The set of observed expressions is augmented by a large set of built-in constructions used to realize common semantic relationships such as "is-a" and "has-a," as well as constructions for the types of messages and propositions the SIGHT system

extracts from the graphics. This enables the generation of novel sentences to build an abstractive summary of the extracted information, albeit with some reused and "canned" expressions. Nearly 80% of human-authored summaries are produced using a cut-and-paste method of re-combining original sentences in new ways [18]. Thus, we view our approach as a roughly analogous process at the surface level (except we actually encode the underlying semantic representation), "cutting" semantically-relevant phrases and "pasting" them together with generalized constructions to generate abstract summaries.

3.4 Rating Content

Once the document analysis phase is finished and the semantic model is complete, the model is then analyzed to discover which pieces of information conveyed in the document are most salient. Intuitively, the entries in the model that contain the most important information, and which are highly connected to other important entries, are the ones that ought to be included in the summary. Several factors[5] are used to determine the importance of information extracted from the document and stored in the semantic model:

1. Completeness of attributes: the percentage of filled-in slots for the PSR (i.e., "saturation level"), and the importance of the entries filling these slots (a recursive value)
2. Number of connections/relationships with other PSRs, and the importance of those entries (a recursive value)
3. Number of expressions realizing the referent in the document text (similar to frequency)
4. Salience based on document structure, rhetorical structure, and importance as assessed by the graph summarization algorithm

3.5 Content Selection

Scoring the model based on these factors produces a set of weights for each entry. These weights are passed along to the graph-based content selection framework developed for the SIGHT system [8], which iteratively selects concepts to be conveyed in the summary according to apriori importance, related and redundant information, and discourse history. Using this approach, we are able to include concept completeness, prevalence, and discourse structure captured by the model weighting, as well as incorporate relationship-based centrality assessment.

3.6 Surface Generation

Once the most salient entries in the semantic model have been selected for inclusion in the summary, the surface generation process begins. The previous

---

[5] Factors 1, 2, and 3 are similar to the dominant slot fillers, connectivity patterns, and frequency criteria proposed by Reimer & Hahn [29].

version of SIGHT [7], which generated descriptions of bar charts only, relied on FUF/SURGE [13] to realize the summaries of graphs in natural language. A large set of templates were used to combine and realize various predicates describing bar charts. However, in order to produce the wider range of constructions necessary to accomodate the article text, and to take advantage of the variety of expressions observed by Sparser and accumulated in the model, the implementation currently in development uses a modern version of Mumble-86 [25] to handle surface realization. For the concepts in the model chosen for inclusion in the summary, we consult the collection of expressions described in Sect. 3.3 and choose from amongst the available options those having the best "fit" (i.e., compatible via substitution or adjunction of TAG trees) enabling these units to be assembled into sentences that are syntactically and semantically valid.

4   Implementation Status

This project is a work in progress and has thus far focused on building the semantic model from text and information graphics. The semantic grammar for Sparser that we have extended is presently capable of producing a nearly-complete parse for several texts in the business news domain (such as the Medtronic article). The SIGHT system is capable of full processing of many simple bar charts (see [10] for limitations), and can identify the intended message in line graphs and grouped bar charts. We are currently working on rating the importance of informational propositions extracted from line graphs, and decomposing these propositions for incorporation into the semantic model. The content rating system remains to be fully implemented and fused with the existing graph-based content selection framework. Finally, a prototype has been developed to use the expressions observed by Sparser for the realization of novel sentences [24], but this component still needs to be integrated with the content rating and selection module. Based on the model built from the Medtronic article, if the resources to be selected by the not-yet-operational content planner are instead chosen by hand, the surface realization component produces the following one-sentence summary:

Wuensch expects a 12-month target of 62 for medical device giant Medtronic.

Company1 ("Medtronic") and Person1 ("Joanne Wuensch," a stock analyst) are the two most prominent concepts in the model (Fig. 2). However, there are no direct links between these concepts, meaning none of the collected phrasings can express them both at the same time. Instead, by using the phrasing provided by a third concept, TargetStockPrice1, we are able to combine all three concepts (via substitution and adjunction operations on the underlying TAG trees), together with a "built-in" realization inherited by the TargetStockPrice category (a subtype of Expectation – not shown in the figure), into the final surface form.

5   Evaluation

Final system evaluation will not be possible until the implementation (in progress) is capable of automatically producing surface output. Summaries generated by

our system will be compared to those of human authors and others created by extractive methods. We will use preference-strength judgment experiments [1] in order to test multiple dimensions of preference (e.g., clarity, completeness). We will also evaluate summaries generated both with and without considering the graphical content, in order to assess the benefits of integrating the contributions of the non-text modalities in the representation of the multimodal document.

6 Future Work

Sparser and KRISP currently require substantial manual effort to build the linguistic and knowledge resources necessary to adapt the system to new domains. Individual grammar rules and ontology definitions must be hand-written by an expert and checked against a corpus of domain texts. Presently, Sparser has decent coverage in the business domain and a few others, but the difficulty of increasing the coverage for broader applications affects scalability. For the implementation currently in development, we are manually extending an existing Sparser grammar on an as-needed basis. While it is relatively trivial to adapt to small changes in an existing domain, adapting to radically-different domains requires a significant amount of resources to be built from the ground-up. In order to automatically adapt the system to new and diverse domains, large-scale learning of additional grammar rules and ontology definitions will be necessary. Promising developments in learning syntactic patterns and ontological relations, as well as machine reading, inspire us to investigate the possibility that these resources may be induced automatically. For example, the Never-Ending Language Learning (NELL) project [5] extracts information from the web in order to extend a structured knowledge base. Similar techniques might be able to build the resources used by our system via automatic domain modeling, with the free-text patterns learned by NELL forming the basis of new Sparser grammar rules.

7 Conclusion

Our approach to automatic summarization of multimodal documents relies on a semantic understanding of text and graphics to construct a unified conceptual model that serves as the basis of generating an abstractive summary. By integrating the knowledge obtained from the graphic with the knowledge obtained from the text at the semantic level, we are able to produce an abstract that treats the entire multimodal document as a single, cohesive message, rather than an assortment of disconnected utterances. This method will generate summaries that are more human-like in nature, while not suffering from coherence and other readability issues related to traditional extractive techniques.

References

1. Belz, A., Kow, E.: Comparing rating scales and preference judgements in language evaluation. In: Proceedings of the 6th International Natural Language Generation Conference. pp. 7–16. INLG 2010, ACL, Trim, Ireland (July 2010)

2. Bhatia, S., Lahiri, S., Mitra, P.: Generating synopses for document-element search. In: Proceeding of the 18th ACM Conference on Information and Knowledge Management. pp. 2003–2006. CIKM '09, ACM, Hong Kong (November 2009)

3. Burns, R., Carberry, S., Elzer, S.: Visual and spatial factors in a bayesian reasoning framework for the recognition of intended messages in grouped bar charts. In: Proceedings of the AAAI Workshop on Visual Representations and Reasoning. pp. 6–13. AAAI, Atlanta (July 2010)

4. Carberry, S., Elzer, S., Demir, S.: Information graphics: an untapped resource for digital libraries. In: Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 581–588. ACM, Seattle (August 2006)

5. Carlson, A., Betteridge, J., Kisiel, B., Settles, B., Jr., E.R.H., Mitchell, T.M.: Toward an architecture for never-ending language learning. In: Proc. of the 24th Conference on Artificial Intelligence. pp. 1306–1313. AAAI, Atlanta (July 2010)

6. Chester, D., Elzer, S.: Getting computers to see information graphics so users do not have to. In: Proceedings of the 15th International Symposium on Methodologies for Intelligent Systems, Lecture Notes in Artificial Intelligence 3488. pp. 660–668. ISMIS 2005, Springer-Verlag, Saratoga Springs, NY (June 2005)

7. Demir, S., Carberry, S., McCoy, K.F.: Generating textual summaries of bar charts. In: Proceedings of the 5th International Natural Language Generation Conference. pp. 7–15. INLG 2008, ACL, Salt Fork, Ohio (2008)

8. Demir, S., Carberry, S., McCoy, K.F.: A discourse-aware graph-based content-selection framework. In: Proceedings of the 6th International Natural Language Generation Conference. pp. 17–26. INLG 2010, ACL, Trim, Ireland (July 2010)

9. Demir, S., Oliver, D., Schwartz, E., Elzer, S., Carberry, S., McCoy, K.F.: Interactive SIGHT into information graphics. In: Proc. of the 2010 Int'l Cross Disciplinary Conference on Web Accessibility. pp. 16:1–16:10. ACM, Raleigh, NC (April 2010)

10. Demir, S., Oliver, D., Schwartz, E., Elzer, S., Carberry, S., McCoy, K.F., Chester, D.: Interactive SIGHT: textual access to simple bar charts. The New Review of Hypermedia and Multimedia 16(3), 245–279 (2010)

11. Demner-Fushman, D., Antani, S., Simpson, M., Thoma, G.R.: Annotation and retrieval of clinically relevant images. International Journal of Medical Informatics 78(12), 59–67 (2009)

12. Dumontier, M., Ferres, L., Villanueva-Rosales, N.: Modeling and querying graphical representations of statistical data. Web Semantics: Science, Services and Agents on the World Wide Web 8(2-3), 241 – 254 (2010)

13. Elhadad, M., Robin, J.: An overview of SURGE: a re-usable comprehensive syntactic realization component. In: Proceedings of the 8th International Natural Language Generation Workshop (Posters & Demos). ACL, Sussex, UK (June 1996)

14. Elzer, S., Carberry, S., Chester, D., Demir, S., Green, N., Zukerman, I., Trnka, K.: Exploring and exploiting the limited utility of captions in recognizing intention in information graphics. In: Proceedings of the 43rd Annual Meeting of the Assn. for Computational Linguistics. pp. 223–230. ACL, Ann Arbor (June 2005)

15. Elzer, S., Carberry, S., Zukerman, I.: The automated understanding of simple bar charts. Artificial Intelligence 175, 526–555 (February 2011)

16. Elzer, S., Schwartz, E., Carberry, S., Chester, D., Demir, S., Wu, P.: Bar charts in popular media: Conveying their message to visually impaired users via speech. In: Ras, Z., Tsay, L.S. (eds.) Advances in Intelligent Information Systems, Studies in Computational Intelligence, vol. 265, pp. 275–298. Springer (2010)

17. Fiszman, M., Rindflesch, T.C., Kilicoglu, H.: Abstraction summarization for managing the biomedical research literature. In: Proceedings of the HLT-NAACL Workshop on Computational Lexical Semantics. pp. 76–83. ACL, Boston (May 2004)
18. Jing, H., McKeown, K.R.: The decomposition of human-written summary sentences. In: Proc. of the 22nd Annual Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval. pp. 129–136. ACM, Berkeley (August 1999)
19. Kupiec, J., Pedersen, J., Chen, F.: A trainable document summarizer. In: Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 68–73. ACM, Seattle (July 1995)
20. Lin, C.Y.: Training a selection function for extraction. In: Proceedings of the 8th International Conference on Information and Knowledge Management. pp. 55–62. CIKM '99, ACM, Kansas City (November 1999)
21. McDonald, D.D.: An efficient chart-based algorithm for partial-parsing of unrestricted texts. In: Proceedings of the 3rd Conference on Applied Natural Language Processing. pp. 193–200. ACL, Trento (March 1992)
22. McDonald, D.D.: Issues in the representation of real texts: the design of KRISP. In: Iwańska, L.M., Shapiro, S.C. (eds.) Natural Language Processing and Knowledge Representation, pp. 77–110. MIT Press, Cambridge, MA (2000)
23. McDonald, D.D.: Partially saturated referents as a source of complexity in semantic interpretation. In: Proceedings of the NAACL-ANLP 2000 Workshop on Syntactic and Semantic Complexity in NLP Systems. pp. 51–58. ACL, Seattle (April 2000)
24. McDonald, D.D., Greenbacker, C.F.: 'If you've heard it, you can say it' - towards an account of expressibility. In: Proceedings of the 6th International Natural Language Generation Conference. pp. 185–190. INLG 2010, ACL, Trim, Ireland (July 2010)
25. Meteer, M., McDonald, D., Anderson, S., Forster, D., Gay, L., Huettner, A., Sibun, P.: Mumble-86: Design and implementation. Technical Report 87-87, Dept. of Computer & Information Science, Univ. of Massachusetts (Sept 1987), 174 pgs.
26. Nenkova, A.: Understanding the process of multi-document summarization: content selection, rewrite and evaluation. Ph.D. thesis, Columbia University (January 2006)
27. Radev, D.R., Hovy, E., McKeown, K.: Introduction to the special issue on summarization. Computational Linguistics 28(4), 399–408 (2002)
28. Rau, L.F., Jacobs, P.S., Zernik, U.: Information extraction and text summarization using linguistic knowledge acquisition. Information Processing & Management 25(4), 419 – 428 (1989)
29. Reimer, U., Hahn, U.: Text condensation as knowledge base abstraction. In: Proceedings of the 4th Conference on Artificial Intelligence Applications. pp. 338–344. CAIA '88, IEEE, San Diego (March 1988)
30. Saggion, H., Lapalme, G.: Generating indicative-informative summaries with SumUM. Computational Linguistics 28, 497–526 (December 2002)
31. Srihari, R.K., Zhang, Z., Rao, A.: Intelligent indexing and semantic retrieval of multimodal documents. Information Retrieval 2(2-3), 245–275 (2000)
32. Wu, P., Carberry, S., Elzer, S., Chester, D.: Recognizing the intended message of line graphs. In: Goel, A., Jamnik, M., Narayanan, N. (eds.) Diagrammatic Representation and Inference, LNCS, vol. 6170, pp. 220–234. Springer (2010)
33. Wu, P., Carberry, S., Elzer, S.: Segmenting line graphs into trends. In: Proceedings of the Twelfth International Conference on Artificial Intelligence. pp. 697–703. ICAI 2010, CSREA Press, Las Vegas (July 2008)
34. Yu, H., Liu, F., Ramesh, B.P.: Automatic figure ranking and user interfacing for intelligent figure search. PLoS ONE 5(10), e12983 (10 2010)