Generating Visual Arguments: a Media-independent Approach

Nancy Green, Stephan Kerpedjiev, Steven F. Roth
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15231 USA
ngreen,kerpedji,roth@cs.cmu.edu

Giuseppe Carenini, Johanna Moore
Intelligent Systems Program
University of Pittsburgh
Pittsburgh, PA 15260 USA
carenini,jmoore@cs.pitt.edu

Introduction

The research reported here is part of our ongoing effort (Kerpedjiev *et al.* 1997b; 1997a; Green, Carenini, & Moore 1998; Green *et al.* 1998; Kerpedjiev *et al.* 1998) to design systems that can automatically generate integrated text and information graphics presentations of complex, quantitative data. In this paper, we take the position that certain types of arguments that can be presented visually in information graphics (e.g., bar charts and scatter plots) can be generated from an underlying media-independent representation of a presentation. In support of this claim, first we briefly describe the architecture we are developing for the generation of integrated text and information graphics presentations. In this architecture, media-independent communicative acts are transformed into user task specifications which are the basis for the automatic design of the presentation's graphics. Then we present an example showing correspondences between the media-independent representation of an argument and the tasks that would be used to design a graphic expressing the argument.

Architecture

We are investigating the integration of two complementary approaches to automatic generation of integrated text and graphics presentations: hierarchical planning to achieve communicative goals and task-based graphic design. Researchers in natural language processing (Mittal *et al.* 1995; Moore 1995; Wahlster *et al.* 1993) have modeled presentation design as a process of hierarchical planning to achieve communicative goals. Researchers in graphics have emphasized the need to design presentations that support the perceptual and logical tasks a user must perform (Casner 1991; Roth & Mattis 1990; Beshers & Feiner 1993).

In our hybrid approach, shown in Figure 1, the content and organization of a presentation is first planned at a media-independent level using a hierarchical planner (Young 1994), resulting in a presentation plan. The presentation plan describes the intentional and informational structure of the presentation (Moser, Moore, & Glendening 1995; Moore & Pollack 1992), as well as what low-level media-independent communicative acts are to be performed by the system to achieve the presentation's goals. A media allocation component decides which parts of the presentation plan to realize in which media. Two media-specific generators (text, graphics) then realize their assigned parts of the plan.

The text generator converts its assigned part of the plan to functional descriptions of sentential units, which are subsequently realized by a general-purpose sentence generator (FUF/SURGE) (Elhadad & Robin 1996). (The complex process of converting the plan to text is beyond the scope of this paper.) Graphics generation is performed in two stages. First, the graphics generator converts the parts of the plan assigned to it by the media-allocation component to a sequence of user tasks that will enable the presentation's goals to be achieved. (Previous integrated text and graphic generation systems, e.g., (Andre & Rist 1994; Fasciano & Lapalme 1996; Feiner & McKeown 1991; Maybury 1991; McKeown *et al.* 1992; Wahlster *et al.* 1993) have not attempted to automatically derive user tasks from a media-independent presentation plan.) The task sequence is then input to the SAGE graphic design system (Roth & Mattis 1990; Roth *et al.* 1994), which automatically creates a graphic designed to enable the user to perform these tasks. All design decisions are made by SAGE, from the type of graphic (e.g., a bar chart), to specific properties of a graphic (e.g., the choice of horizontal as opposed to vertical bars). In this way, graphic design is tailored to a presentation's goals. (For details on our approach to graphics generation, including the derivation of tasks in our system, see (Kerpedjiev *et al.* 1998).)

Expressing an argument in graphics

In this section, we give an analysis of an argument and its representation in a media-independent presentation plan. We describe the user tasks which would be derived from the media-independent communicative acts of the plan in our current approach, and then suggest some ways in which the structure of the discourse may also contribute to the design of effective graphics, as

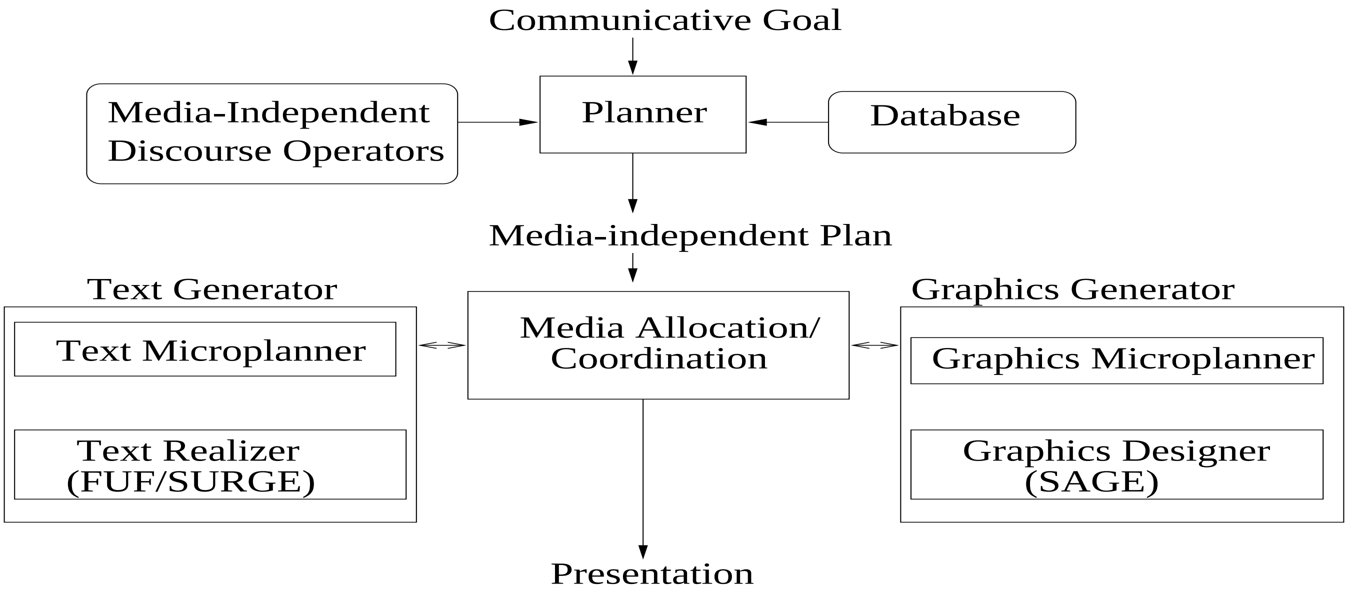

Figure 1: Integrated Text-Graphics Generation Architecture

well as its influence on media allocation.

Analysis of example

The goal of the example presentation[1] is for the user to accept the belief that a certain local newspaper, the Post-Gazette (PPG), has more readers than the total number of readers of all other newspapers that are subscribed to in some region. The user is currently ignorant of this fact, but probably would not accept it just on the basis of a simple assertion by the system, due to his current beliefs. In particular, the user knows that the New York Times (NYT) has more readers than the Wall Street Journal (WSJ) in the region, and erroneously believes that because of this, the New York Times must have the largest number of readers of all newspapers in the region. However, the latter belief is incompatible with the belief which it is the goal of the presentation to get the user to accept.

To simplify discussion, let us abbreviate the propositions playing a role in the example as follows:

- **Q**: the number of readers of NYT exceeds the number of readers of WSJ
- **R**: the number of readers of NYT exceeds the number of readers of any other paper in the region
- **T**: the number of readers of PPG exceeds the total number of readers of all other papers in the region.

An argument strategy for this situation can be defined as follows. If the stated constraints hold, and the stated subgoals are achieved, then the stated goal will be achieved:

[1]The data used in this example is fictitious.

- Goal: **(BMB User T)**, i.e., the goal is that the User believe that it is mutually believed with the system that T,[2]
- Constraints:
  - **(Bel System (Bel User Q))**, the system believes that the user believes that Q,
  - **(Bel System (Bel User (Entails Q R)))**, the system believes that the user believes that Q entails R.
  - **(Bel System (Incompatible R T))**, the system believes that R and T cannot both be true,
  - **(Bel System Q)**, the system believes that Q,
  - **(Bel System T)**, the system believes that T, and
  - **(Bel System (not R))**, the system believes that R is false,
- Subgoals:
  - **(BMB User Q)**, i.e., make sure that the User knows that the System is aware of Q,
  - **(BMB User (not (Entails Q R)))**
  - **(BMB User (Bel System T))**

One way in which each of the above three subgoals could be achieved is by the following actions:

- **(Acknowledge System User Q)**, which achieves **(BMB User Q)**.
- **(Assert System User (not R))**, which achieves **(BMB User (not (Entails Q R)))** provided that **(BMB User Q)**.

[2]*BMB* is a one-sided mutual belief operator (Clark & Marshall 1981); *(BMB agent proposition)* is interpreted as *agent believes it to be mutually believed with the other agent that proposition holds.*

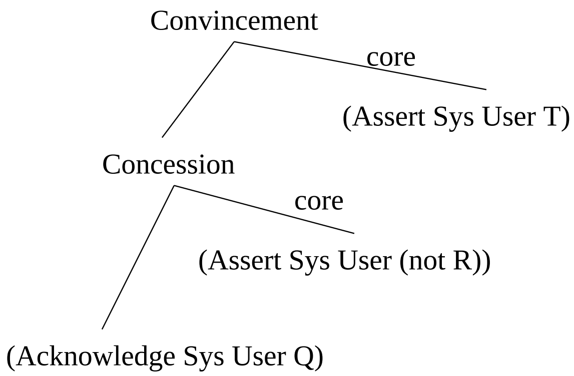

Figure 2: Discourse Plan

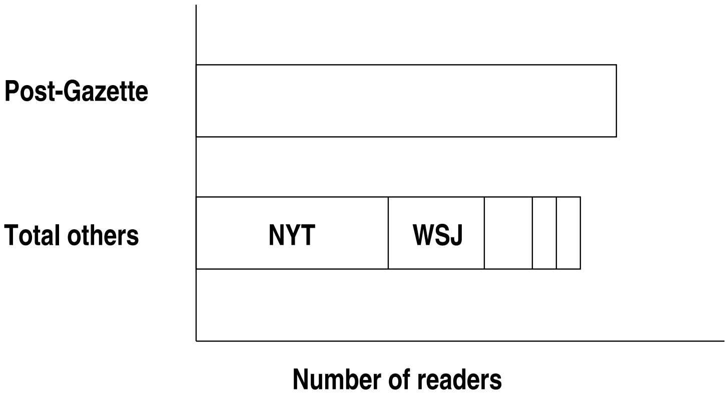

Figure 3: Graphic realizing the argument

- **(Assert System User T)**, which achieves **(BMB User (Bel System T))**.

In summary, the intentional structure of this argument can be represented in the plan shown in Figure 2. (The figure shows only the hierarchical relations among the communicative acts and the core-contributor distinction among acts; unlabelled acts are contributors.) Such a plan might be realized in text as follows: *Although the New York Times is read by more people in Western PA than the Wall Street Journal, the New York Times does not have the highest number of readers in the region. The Post-Gazette has more readers than the total number of readers of all other newspapers in the five-county Western PA region.*

Realization in Graphics

A graphic realizing this argument is shown in Figure 3. The core of the argument, the assertion that T holds, is expressed by enabling the user to perform the task of comparing the upper bar's length (which represents the number of readers of PPG) to the lower bar's length (which represents the total number of readers of the other newspapers). (In general, an assertion that some quantity is greater than another quantity would be transformed by our graphics generator into a comparison task; for details on the process of deriving a task sequence, see (Kerpedjiev *et al.* 1998).)

The contributor given in support of T is expressed in the same graphic, although less prominently. Its core consists of the assertion (not R), which is expressed in the graphic indirectly by *falsifying* R. That is, by enabling the user to perform the task of comparing the length of the segment of the lower bar labelled NYT to the length of the upper bar, the user can see that there is one newspaper (PPG) with more readers than NYT, which falsifies R. The concession Q (contributing to the acceptance of the assertion that R does not hold), is expressed within the lower bar by enabling the user to perform the task of comparing the length of the segment labelled NYT (representing the number of NYT readers) to the length of the segment labelled WSJ (representing the number of WSJ readers).

Thus, in our current approach each the three low-level communicative acts of this plan would be transformed into the comparison tasks described above. The tasks would then be used by SAGE to design a graphic such as the one shown in Figure 3 to support these tasks. Note that more than one graphic may be designed by SAGE to support the tasks. An interesting open question that we are investigating is how the intentional structure of an argument should influence graphic design. In this graphic, for example, the assertion corresponding to the core of the argument is more visually prominent than the other information since the graphic contains only two horizontal bars, one for each of the entities compared in the core assertion. Also, the quantities compared in the core assertion are encoded differently (i.e. as horizontal bars) from the other quantities (i.e. as segments of a stacked bar). The user may interpret the difference as conversationally implicating an important distinction in the two sets of quantities (Marks & Reiter 1990).

A related issue is the role of discourse structure in media allocation. For example, if the graphic shown in Figure 3 is accompanied only by text realizing the core of the argument (e.g., *The Post-Gazette has more readers than the total number of readers of all other newspapers in the five-county Western PA region*), then the text would contribute to the user's recognition of the main point of the graphic. On the other hand, if the same graphic is accompanied only by text realizing one of the other acts of the plan (e.g., *The New York Times has more readers than the Wall Street Journal*), then the text might impede the user's recognition of the main point of the graphic. In future work, we hope to address these open issues.

Acknowledgments

This project was supported by DARPA, contract DAA-1593K0005.

References

Andre, E., and Rist, T. 1994. Referring to world objects with text and pictures. In *Proceedings of COLING*, 530–34.

Beshers, C., and Feiner, S. 1993. AutoVisual: Rule-based design of interactive multivariate visualization. *IEEE Computer Graphics and Applications* 41–49.

Casner, S. M. 1991. A task-analytic approach to the automated design of information graphic presentations. *ACM Transactions on Graphics* 10(2):111–151.

Clark, H., and Marshall, C. 1981. Definite reference and mutual knowledge. In Joshi, A. K.; Webber, B.; and Sag, I., eds., *Elements of discourse understanding*. Cambridge: Cambridge University Press.

Elhadad, M., and Robin, J. 1996. An overview of SURGE: A reusable comprehensive syntactic realization component. Technical Report Technical Report 96-03, Dept of Mathematics and Computer Science, Ben Gurion University, Beer Sheva, Israel.

Fasciano, M., and Lapalme, G. 1996. PostGraphe: a System for the Generation of Statistical Graphics and Text. In *Proceedings of the 8th International Natural Language Generation Workshop*, 51–60.

Feiner, S., and McKeown, K. 1991. Automating the generation of coordinated multimedia explanations. *IEEE Computer* 24(10):33–40.

Green, N.; Carenini, G.; Kerpedjiev, S.; Roth, S.; and Moore, J. 1998. A media-independent content language for integrated text and graphics generation. Submitted.

Green, N.; Carenini, G.; and Moore, J. 1998. Generating attributive descriptions for integrated text and information graphics presentations. In *Proceedings of the Ninth International Workshop on Natural Language Generation*. To appear.

Kerpedjiev, S.; Carenini, G.; Roth, S.; and Moore, J. 1997a. AutoBrief: a multimedia presentation system for assisting data analysis. *Computer Standards and Interfaces* 18:583–593.

Kerpedjiev, S.; Carenini, G.; Roth, S. F.; and Moore, J. D. 1997b. Integrating planning and task-based design for multimedia presentation. In *International Conference on Intelligent User Interfaces (IUI '97)*, 145–152. Association for Computing Machinery.

Kerpedjiev, S.; Carenini, G.; Green, N.; Roth, S. F.; and Moore, J. D. 1998. Saying it in graphics: from intentions to visualizations. Submitted.

Marks, J., and Reiter, E. 1990. Avoiding unwanted conversational implicatures in text and graphics. In *Proceedings of the National Conference on Artificial Intelligence*, 450–455.

Maybury, M. T. 1991. Planning multimedia explanations using communicative acts. In *Proceedings of the Ninth National Conference on Artificial Intelligence*, 61–66.

McKeown, K.; Feiner, S.; Robin, J.; Seligmann, D.; and Tanenblatt, M. 1992. Generating cross-references for multimedia explanation. In *Proceedings of AAAI*, 9–16.

Mittal, V.; Roth, S.; Moore, J.; Mattis, J.; and Carenini, G. 1995. Generating explanatory captions for information graphics. In *Proceedings*, 1276–1283. Montreal: International Joint Conference on Artificial Intelligence.

Moore, J. D., and Pollack, M. E. 1992. A problem for RST: The need for multi-level discourse analysis. *Computational Linguistics* 18(4):537–544.

Moore, J. D. 1995. *Participating in Explanatory Dialogues*. MIT Press.

Moser, M.; Moore, J. D.; and Glendening, E. 1995. Instructions for coding explanations: Identifying segments, relations and minimal units. Technical report, Learning Research and Development Center, University of Pittsburgh, Pittsburgh, Pennsylvania.

Roth, S., and Mattis, J. 1990. Data characterization for intelligent graphics presentation. In *Proceedings of the Conference on Human Factors in Computing Systems (SIGCHI '90)*, 193–200.

Roth, S. F.; Kolojejchick, J.; Mattis, J.; and Goldstein, J. 1994. Interactive graphic design using automatic presentation knowledge. In *Proceedings of the Conference on Human Factors in Computing Systems (SIGCHI '94)*, 112–117.

Wahlster, W.; Andre, E.; Finkler, W.; Profitlich, H.-J.; and Rist, T. 1993. Plan-based integration of natural language and graphics generation. *Artificial Intelligence* 63:387–427.

Young, M. R. 1994. A developer's guide to the longbow discourse planning system. Technical Report ISP TR Number: 94-4, University of Pittsburgh, Intelligent Systems Program.