

# Answering it with charts — Dialogue in natural language and charts —

**Tsuneaki KATO**

Graduate School of Arts and Sciences, The University of Tokyo  
3-8-1 Komaba, Meguroku, Tokyo 153-8902 Japan  
kato@boz.c.u-tokyo.ac.jp

**Mitsunori MATSUSHITA and Eisaku MAEDA**

NTT Communication Science Laboratories, NTT Corp.  
2-4 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237 Japan  
{mat, maeda}@cslab.kecl.ntt.co.jp

## Abstract

A methodology is proposed for taking queries and requests expressed in natural language as input and answering them in charts through organizing that interaction into felicitous dialogue. Charts and graphics, as well as languages, are important modes of communication. This is especially true of those which are used frequently when people analyze huge amount of data interactively, in order to find out its characteristics or to resolve questions about it. This paper raises the problem that in such situations the correctness of the charts depends on the context, and proposes a framework to resolve it. The core of the framework is a logical form that includes the specifications of the user's perspective and the proper treatment of the logical form for handling utterance fragments. The framework has been implemented and confirmed to be appropriate.

## 1 Introduction

Charts and graphics, as well as languages, are important modes of communication. Considering this importance, the automatic design of charts and graphics suitable for achieving a given communicative purpose has been studied actively (Maybury and Wahlster, 1998). It has been demonstrated that the characteristics of data drawn on the chart (Mackinlay, 1986), the intention to be achieved through the chart (Roth and Mattis, 1990), and the task accomplished using the chart (Casner, 1991) play important roles in designing appropriate charts. The automatic design of multimedia documents in which those charts coordinate with natural language text has also been studied (Roth et al., 1991; Kerpedjiev et al., 1998).

In that research, systems take an assertion to be conveyed or a communicative goal to be achieved, and design the most appropriate multimedia pre-

sentation for that purpose. The purpose of drawing charts and graphics, however, is not restricted just to use in such presentations. In particular, as the drawing of quantitative charts helps to analyze huge amount of data and to find out its characteristics, they can be a useful means for interactive exploratory data analysis. An analyst, led by an interest or question, draws a chart, then a new interest or question comes up and she draws another chart. Through this process, the analyst finds out a characteristic behind the data or understands the reason for it.

This paper discusses automatic quantitative chart design that supports such interactive exploratory data analysis. That is, a methodology is proposed for taking queries and requests expressed in natural language as input and answering them in charts through organizing that interaction into felicitous dialogue. The objectives of our research are an automatic chart design that considers dialogue context, and a dialogue mechanism that uses charts as its output mode.

## 2 Context sensibility of chart appropriateness

Let us suppose an analyst, planning sales of her company's products, gets interested in its sales in a particular district. So she requests the following from a system in front of her:

- (1) Show me the sales in the Shikoku district for '93 and '94.

The system answers this request by drawing an appropriate chart. She continues and utters a fragment of a request:

- (2) By prefecture.

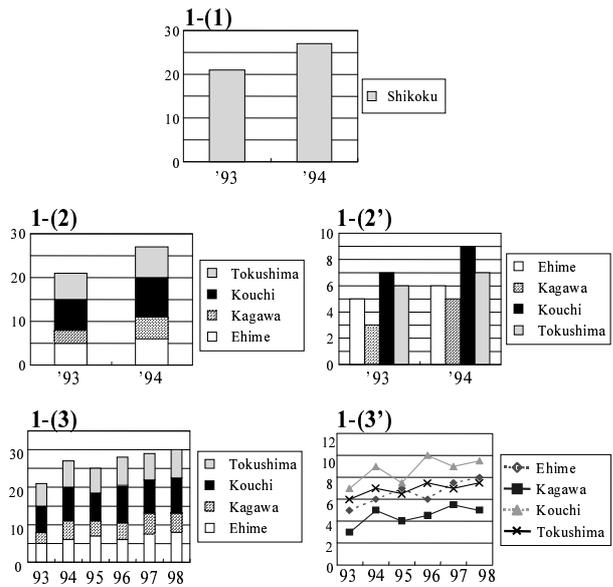


Figure 1: Examples of Context Sensibility of Chart Correctness (1)

The system understands this fragment and draws a new chart. This dialogue might be followed by an utterance like:

(3) Through '98.

The correct chart made in the response to utterance (1) is like that shown in Figure 1-(1). Following this chart, the response to utterance (2) should be made by the stacked column chart, as shown in Figure 1-(2), in which each column is subdivided in order to show the sales by prefecture. Here, these two consecutive utterances have a combined meaning similar to:

(2') Show me the sales in each prefecture of the Shikoku district for '93 and '94.

It is interesting that when utterance (2') is given without a specific context, the response should be made by the grouped column chart, as shown in Figure 1-(2'), rather than as in Figure 1-(2). The preference for Figure 1-(2) as the response to utterance (2) may come from the perspective represented in utterance (1) that she wants to look at the sales of the district in total or the similarity in shape between Figure 1-(2) and 1-(1). In any case, it is important that an appropriate chart form depends on what utterance or series of utterances was used to express a request and what chart has been drawn previously. This context sensibility of chart appropriateness oc-

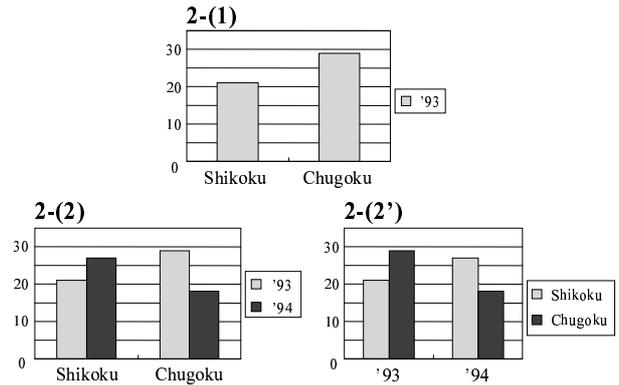


Figure 2: Examples of Context Sensibility of Chart Correctness (2)

curs constantly. For utterance (3) in the previous dialogue, the chart shown in Figure 1-(3) is preferred when Figure 1-(2) is used to answer the previous request. The chart of Figure 1-(3') is preferred, however, as the response to utterance (3'), which combines requests (1), (2) and (3) into one;

(3') Show me the sales in each prefecture of Shikoku district from '93 through '98.

The chart form is not the only dimension sensitive to dialogue context. Consider the following example.

(4) Show me the sales in Shikoku and Chugoku for '93.

(5) Add the one for '94.

(5') Show me the sales in Shikoku and Chugoku for '93 and '94.

The preferred response to utterance (5) must be in the chart shown in Figure 2-(2) when preceding utterance (4) was answered by the chart shown in Figure 2-(1), while the chart of Figure 2-(2') would be used when the request is just utterance (5'). It is clear that the decision on axis assignment is also context sensitive.

In conventional ellipsis handling (Hendrix et al., 1978; Carbonell and Hayes, 1983), the interpretation of an utterance fragment, such as utterance (2) following utterance (1), is the same as the interpretation of utterance (2'). When a response based on this interpretation is made, the data plotted on the chart may be correct, but the chart form and style cannot be. This implies that something extra is needed for handling dialogue in charts, and it

is insufficient just to combine two mechanisms for non-interactive automatic chart design and natural language dialogue understanding.

### 3 Handling dialogue in natural language and charts

This section proposes a methodology for handling dialogue in natural language and charts. First, a logical form that represents the interpretation of utterances is proposed. Then, how to represent the perspectives from which the user wants to look at the data and how to relate them to chart realization are described. Last, a way of handling utterance fragments is discussed.

#### 3.1 A logical form

A logical form that describes the content of a given utterance must be able to represent what region of data the user is interested in and what type of analysis she wants to conduct in order to obtain the data to be plotted. Moreover, it must convey the intention of the user, that is, what information she wants to obtain through that chart. The logical form we propose, which satisfies those requirements, takes the form,  $\{Description, \dots; Action\}$ , where *Action* specifies the main speech act in a given utterance, and *Descriptions* describe constraints or conditions that the objects related to the action should satisfy. *Action* can be a request to display charts or a request for information conveyed through charts. In this paper, however, it only covers requests to display charts, which takes the form `reqDisplay(ListofVars, ListofAspects)`, where *ListofVars* is the list of variables plotted on the chart. *ListofAspects* is the list of aspects of the data the user is focusing on and represents the perspectives from which she wants to look at the data. *Descriptions* describe constraints or conditions that the objects related to the action should satisfy, which has the form,  $\{Quantifier, Var/Class, Restriction\}$ , where *Quantifier* is a generalized quantifier, *Var* is the variable of quantification, and the quantification ranges over the objects each of which is a member of *Class* and satisfies *Restriction*. That is, this logical form is a flattened version of Woods' MRL (Woods, 1978), and as in Woods' MRL, *Class* can be a function. Moreover, *Classes*, each of which each variable and object is associated with, are hierarchically organized and represent not only the domain an object is classified into, but also its granularity. An object that belongs to the *area* domain,

for example, belongs to one of classes: *district*, *prefecture*, or *city*, according to its granularity. The subsumption relation is defined between objects that belong to classes with different granularities and the same domain. In *Restriction*, implicit coercion between granularities is allowed and aggregation such as summation is represented implicitly using this mechanism.<sup>1</sup>

For example, utterance (2') is interpreted into the logical form:

```
{every, X/year, member(X, [1993/year, 1994/year])},
 {every, Y/prefecture, Y=shikoku/district},
 {the, Z/sales(X, Y), true};
 reqDisplay([Z], [hCompBy(X), dCompBy(Y)])}
```

The first description states that variable *X* ranges over two objects of year class, 1993 and 1994. The second description states variable *Y* ranges over objects with *prefecture* granularity that are subsumed by *Shikoku*, which is itself an object with *district* granularity. In this case, the equality in the restriction coerces into the subsumption relation. In the third description, `sales(X, Y)` is a function from time and area to sales amounts. The perspective is discussed in the next section.

#### 3.2 Perspectives and chart realization

The second argument of the `reqDisplay` action specifies the user's perspective, which is the list of aspects of the data she is focusing on. The following basic aspects are covered for the present. Suppose *Var1* is an element of the first argument of the action, that is, a variable to be plotted. The `compBy(Var2)` specifies that the user is interested in comparison of the values of *Var1* by each possible *Var2* instance. In other words, she is focusing on the difference in the values of *Var1* which came from the difference of *Var2* instances. The `totalIn(Var2)` specifies that the user also wants to know the total for the values of *Var1* in all *Var2* instances. The `propTo(Var2)` specifies that she is interested in the proportion of each value of *Var1* to the total in all *Var2* instances. The `changeWith(Var2)` specifies that she is interested in the change of value of *Var1* with the change of *Var2* instances. The `compBy(Var2)` is, in a sense, a default since the quantitative chart, in principle, shows variation of the values of a dependent variable by the difference of the values of independent variables.

<sup>1</sup>The way to create a data table to be plotted on a chart from raw material data using this type of logical form was discussed in (Matsushita et al., 2000).

From now on, our discussion is restricted to the two-dimensional chart forms for up to two independent variables and one dependent variable. These forms of charts are widely used and still have plenty of variety. In chart realization under this restriction, one of the independent variables, which are the arguments of the function in the logical form, is assigned to the horizontal axis and the other to the hidden or depth axis. For example, variable  $X$  is assigned to the horizontal axis and variable  $Y$  is assigned to the depth axis in the realization of the logical form of the previous example into Figure 1-(2'). One of the realization parameters is this axis assignment of variables.

Let us think about how the perspective guides chart realization. In two-dimensional charts, changes of values can be displayed along the horizontal axis using, say, the line chart, while total and proportion is easy to show on a depth axis by using, say, the stacked chart. Comparison can be exhibited on either axis. Therefore, `compBy` and `changeWith` are possible for the variable assigned to the horizontal axis, and `compBy`, `totalIn`, and `propTo` are possible for the variable assigned to the depth axis. Next, let us think about what combinations of aspects are possible and sufficient for the perspective. Since we cannot compare the difference nor see the changes regarding a variable that has only one instance in the range specified by the restriction, the aspect for such a variable, which we call a uniquely instantiated variable, is meaningless. Excluding that variable, each remaining variable takes only one aspect. Their combination should result in those aspects being properly assigned to the axes. Here, by subcategorizing `compBy`, which can be assigned to either axis, into `hCompBy` and `dCompBy`, which should be assigned to the horizontal and depth axis respectively, you can assign an axis to the variables only by assigning one aspect to each. Thus, when two independent variables,  $X$  and  $Y$ , are not uniquely instantiated, one of them, say  $X$ , takes either `hCompBy(X)` or `changeWith(X)`, and the other, say  $Y$ , takes either `dCompBy(Y)`, `totalIn(Y)`, or `propTo(Y)`. In this case, variable  $X$  is assigned to the horizontal axis and variable  $Y$  to the depth axis. Both variables are not allowed to take aspects assigned to the same axis. The combination of `dCompBy(X)` and `totalIn(Y)` is forbidden, for example. Figure 3 summarizes the relationship between combinations of aspects and realized chart forms.

Now, chart realization is reduced to aspect assignment. In other words, you can realize an appropriate chart from a given user utterance by obtaining its logical form, and, if some of the aspects are implicit in the utterance, by completing the perspective using the characteristics of the data concerned, the current context and so on.<sup>2</sup> Obtaining and completing the perspective, which is vital for chart realization, is carried out as follows.<sup>3</sup>

- A portion of the perspective is explicit in the utterance. For example, it is obvious that the user is focusing on the change over time when she says that she wants to see the change in sales. Such perspectives are acquired through the interpretation of her utterance.
- Characteristics of statistics returning a dependent variable sometimes suggest the perspective. For example, share and profitability suggest her interest in the proportion to the total. Such characteristics are used for completing the perspective.
- The class of independent variables also suggests the perspective. A class whose instances constitute an interval scale, such as a time series, suggests changes over it (i.e. `changeWith`) for its aspect to be focused on when the variable ranges over a lot of instances. Even when it has a few instances, comparisons by it (i.e. `hCompBy`) are preferred. This criterion for selecting between `changeWith` and `hCompBy` is used throughout this paper.
- When variables with more than one instance are left with no aspect, a default is used. When just one such variables is left, it takes `changeWith` or `hCompBy` if possible, otherwise it takes `dCompBy`. When two such vari-

---

<sup>2</sup>Chart realization has dimensions other than those discussed here. For example, while the independent variable is always assigned to the vertical axis in our discussion, it can be assigned to the horizontal axis. The rank of instances on an axis, the scales of axes, and visual prompts such as labels and arrows are also dimensions which should be considered (Mital, 1998; Fasciano and Lapalme, 1996). Although discussion of those dimensions exceeds the scope of this paper, we believe that a natural extension of perspective would cover them.

<sup>3</sup>In implementation, perspective completion based on these features is more sophisticated, referring to heuristics from a textbook for drawing charts (Zelazny, 1996) and knowledge acquired from a chart corpus using a machine learning technique (Yonezawa et al., 2000).

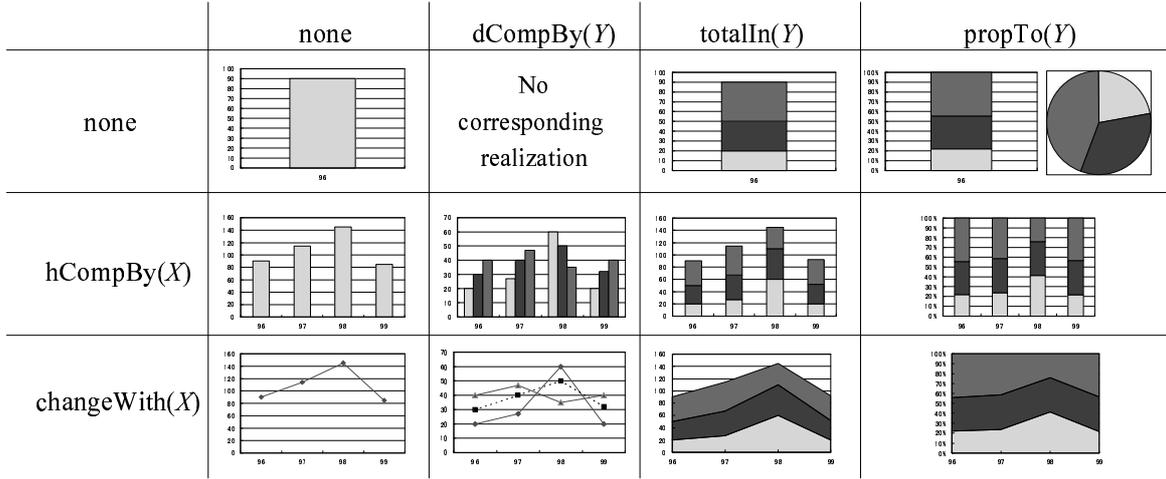


Figure 3: Relationship between Perspectives and Chart Forms

ables are left, one takes hCompBy, the other takes dCompBy.

### 3.3 Utterance fragments and chart appropriateness

Utterance fragments in a specific context should be interpreted not as a logical form, but as a request to revise the logical form given as the context. In order to correctly handle chart appropriateness, which is sensitive to the context, the perspective in the logical form should be revised appropriately according to that request. First, utterance fragments are classified into the following categories according to what part of the logical form should be revised as a result of those fragments. Examples are shown, which are assumed to have been uttered after the utterance, “Show me the sales in Shikoku and Chugoku from ’93 through ’95.”

**Domain alteration:** The *Restriction* part of an independent variable is revised. Examples, “Just Shikoku.”, “Through ’98.”

**Granularity alteration:** The *Class* part of an independent variable is revised to one with a different granularity. Examples, “By quarter.”, “By prefecture.”

**Statistic alteration:** The function returning the dependent variable, which locates its *Class* part, is revised. Examples, “Show me the number of the dealers.”

**Perspective alteration:** The perspective is revised. Examples, “Show me the change.”, “How about the total?”

The type of content words, clue words, and specific phrases contained are exploited for interpreting utterance fragments and for classifying them into one of the above categories. Using those, we can identify what part of the logical form should be changed and how. In addition to the revisions identified, appropriate revision of the perspective is needed for correct chart realization. Revisions of perspective are summarized as follows.<sup>4</sup>

- As a result of domain alteration on variable  $X$ , if the number of instances of  $X$  turns into more than one and the current perspective includes no aspect relating to  $X$ , that is,  $X$  is a uniquely instantiated variable, check `changeWith( $X$ )`, `hCompBy( $X$ )`, and `dCompBy( $X$ )` in this order, and add the first possible one to the perspective. On the other hand, if  $X$  turns into a uniquely instantiated variable, delete the aspect related to  $X$  from the perspective.
- As a result of granularity alteration on variable  $X$ , if the number of instances of  $X$  turns into more than one and the current perspective includes no aspect related to  $X$ , check `changeWith( $X$ )`, `hCompBy( $X$ )`, `totalIn( $X$ )`, and `dCompBy( $X$ )` in this order, and add the first possible one to the perspective. Here, `totalIn( $X$ )` is possible only when the statistics concerned use summation for aggregation. On the other hand, if  $X$  turns into

<sup>4</sup>Revisions of perspective for statistic alterations and perspective alterations are omitted, because the space is limited and our concern is to trace the changes of the user’s perspective especially when she does not mention them explicitly.

a uniquely instantiated variable, delete the aspect related to  $X$  from the perspective.

#### 4 Examples

This section demonstrates how our proposal addresses the problems raised. First, let us consider series of utterance (1), (2) and (3). The interpretation of utterance (1) is

```
{every,X/year,member(X,[1993/year,1994/year])},
{every,Y/district,Y=shikoku/district},
{the,Z/sales(X,Y),true};
reqDisplay([Z],[hCompBy(X)])}
```

As for the perspective, variable  $X$ , which represents time series and ranges over only two instances, obtains not `changeWith(X)`, but `hCompBy(X)`. Variable  $Y$ , which ranges over areas, is uniquely instantiated and no aspect is given to it. The chart realized from this logical form is the column chart of Figure 1-(1). The chart form is determined from the perspective by referring to Figure 3. When utterance (2) is given in this context, it is interpreted as a granularity alteration on variable  $Y$ , and `{every,Y/district,Y=shikoku/district}` of the logical form is revised to `{every,Y/prefecture,Y=shikoku/district}`. In addition, `totalIn(Y)` is added to the perspective, since no aspect related to  $Y$  was in it and sales is a statistic for which summation is used for aggregation. The perspective ends up with `[hCompBy(X), totalIn(Y)]`, and the new chart is the stacked column chart of Figure 1-(2). Utterance (3) in this context is interpreted as a domain alteration on variable  $X$ . Since the perspective remains the same as before, while the restriction of variable  $X$  is revised according to the utterance, the chart obtained is of the same form, which is shown in Figure 1-(3).

On the other hand, for utterance (2'), as neither its expression nor its statistic implies a specific aspect, the perspective is determined according to the characteristics of the independent variables. First, variable  $X$  representing time series obtains `hCompBy(X)`, and then variable  $Y$  ranging over areas is given `dCompBy(Y)`. The perspective ends up with `[hCompBy(X), dCompBy(Y)]`, and the chart realized is the grouped column chart shown in Figure 1-(2'). For utterance (3'), almost the same criteria are applied, but the aspect given to variable  $X$  is `changeWith(X)` as it has many instances. As a result, the grouped line chart of Figure 1-(3') is realized.

Let us move to series of utterance (4) and (5). For utterance (4), since variable  $X$  representing time series is uniquely instantiated, no aspect is given. Then variable  $Y$  ranging over districts can obtain `hCompBy(Y)`. By utterance (5) following it, domain alteration on  $X$  is specified, and  $X$  obtains the possible aspect `dCompBy(X)`, since `hCompBy(Y)` already exists. On the other hand, for utterance (5'), variable  $X$  representing time series precedes variable  $Y$ , obtaining `hCompBy(X)`. As a result, Figure 2-(2), the response to utterance (5), and Figure 2-(2'), the response to utterance (5'), are different in their axis assignments.

#### 5 Discussion

Appropriateness of charts is known to be a function of several factors. This paper revealed that discourse context is one of those factors in addition to those already known such as characteristic of data and user intentions, and proposed a methodology for addressing that factor properly. To our knowledge, there are few studies on automatic chart design for such interactive situations as discussed here. There are many studies on interactive graph drawing of course (Roth et al., 1994). Those, however, are concerned with tools for producing a graph interactively that achieves the user's intention. Their standpoint differs from ours, and the mode used for their interactions is direct manipulation not natural language.

Our proposal partially overlaps with recent studies on automatic chart design. Our logical form has a lot in common with the content language in (Green et al., 1998). The objective of their research, however, is to describe communicative goals to be achieved through generating graphics and text, and differs from ours, which is to describe the user's requests in order to respond to them using charts. Our perspective plays a similar role to that of intention in PostGrappe (Fasciano and Lapalme, 1996). However, there is a crucial difference in that, while their intention is given as input, our perspective is acquired from the user's utterances, data characteristics and dialogue context.

Most of the framework proposed in this paper has been implemented. The prototype system accepts such a wide range of Japanese expressions that the appropriateness of our proposal can be confirmed, though the understanding of those expressions is driven by a simple pattern-based mechanism. The matter worth special mention is the pro-

prototype's chart redraw mechanism. As suggested in this paper, a correct chart in a certain context is one that can be realized by minimum change to the chart as the context, and loses minimum information from it. We supposed that changes and loss of information mattered because considerable mental loads are needed to relate new information in the new chart to old information in the context. In order to reduce those loads, we made the process of change visually understandable. Specifically, the system shows animations that represent which component of the current chart moves and changes to which component of the new one. This feature has received a favorable reception in demonstrations.

A lot remains for future work. First, we will consider answering wh- and yes-no interrogatives in charts in addition to answering requests to show charts. Answering such questions fluently requires collaboration between charts and text. Then, we will examine richer chart realization. We should consider not only increasing the kinds of chart form covered, but also assembling more than one chart to achieve a certain goal. In addition, dimensions such as visual prompts should be incorporated. Lastly, in the context of information visualization, the effects of animation introduced in the implementation of the prototype should be measured quantitatively to prove that it really can reduce mental load rather than merely make itself conspicuous.

## References

- Jaime G. Carbonell and Phillip J. Hayes. 1983. Recovery strategies for parsing extragrammatical language. *American Journal of Computational Linguistics*, 9(3-4):123–146.
- Stephen M. Casner. 1991. A task-analytic approach to the automated design of graphic presentations. *ACM Trans. on Graphics*, 10(2):111–151.
- Massimo Fasciano and Guy Lapalme. 1996. Post-graphic: A system for the generation of statistical graphics and text. In *Proc. 8th International Workshop on Natural Language Generation*, pages 51–60.
- Nancy Green, Giuseppe Carenini, Stephan Kerpedjiev, Steven Roth, and Johanna Moore. 1998. A media-independent content language for integrated text and graphics generation. In *Proc. Workshop on Content Visualization and Intermedia Representations*, pages 69–75.
- Gary G. Hendrix, Earl D. Sacerdoti, Daniel Sgalowicz, and Jonathan Slocum. 1978. Developing a natural language interface to complex data. *ACM Trans. on Database*, 3(2):105–147.
- Stephan Kerpedjiev, Giusappe Carenini, Nancy Green, Johanna Moore, and Steven F. Roth. 1998. Saying it in graphics: from intentions to visualizations. In *Proc. of the Symposium on Information Visualization*, pages 97–101.
- Jock Mackinlay. 1986. Automating the design of graphical presentations of relational information. *ACM Trans. on Graphics*, 5(2):110–141.
- Mitsunori Matsushita, Hayato Yonezawa, and Tsuneaki Kato. 2000. A frame representation of user requirements for automated data visualization. In *Proc. 14th European Conference on Artificial Intelligence*, pages 631–635.
- Mark T. Maybury and Wolfgang Wahlster, editors. 1998. *Readings in Intelligent User Interface*. Morgan Kaufmann Publishers.
- Vibhu O. Mittal. 1998. Visual prompts and graphical design: A framework for exploring the design space of 2-d charts and graphs. In *Proc. 14th National Conference on Artificial Intelligence*, pages 57–63.
- Steven F. Roth and Joe Mattis. 1990. Data characterization for intelligent graphic presentation. In *Proc. Conference on Human Factors in Computing Systems*, pages 193–200.
- Steven F. Roth, Joe Mattis, and Xavier Mesnard. 1991. Graphics and natural language as components of automatic explanation. In J. W. Sullivan and S. W. Tyler, editors, *Intelligent User Interfaces*. ACM Press.
- Steven F. Roth, John Kolojchick, Joe Mattis, and Jade Goldstein. 1994. Interactive graphics design using automatic presentation knowledge. In *Proc. Conference on Human Factors in Computing Systems*, pages 112–117.
- William A. Woods. 1978. Semantics and quantification in natural language question answering. In M. Yovits, editor, *Advances in Computers*, volume 17. Academic Press.
- Hayato Yonezawa, Mitsunori Matsushita, and Tsuneaki Kato. 2000. Criteria to choose appropriate graph-types. In *Proc. Pacific Rim International Conference on Artificial Intelligence*, page 823.
- Gene Zelazny. 1996. *Say It with Charts*. McGraw-Hill, third edition.