

ACCESSIBLE BAR CHARTS FOR VISUALLY IMPAIRED USERS

Stephanie Elzer and Edward Schwartz
Department of Computer Science
Millersville University
Millersville, PA USA
email: elzer@cs.millersville.edu

Sandra Carberry, Daniel Chester, Seniz Demir, and Peng Wu
Department of Computer and Information Sciences
University of Delaware
Newark, DE USA
email: {carberry,chester,demir,pwu}@cis.udel.edu

ABSTRACT

This paper presents a novel approach to enabling visually impaired users to gain access to bar charts on the Web. Our approach differs from previous work by providing the user with the message and knowledge that one would gain from viewing the graphic rather than providing alternative access to the appearance of the graphic. The user interface to the system is implemented as a browser extension. The output of the system is a textual summary, the core content of which is the hypothesized intended message of the graphic designer, as inferred by our Bayesian network. The summary is conveyed to the user by screen reading software. User evaluations have shown the system to be both useful and effective.

KEY WORDS

assistive technology, artificial intelligence, information graphics, natural language processing

1 Introduction

Information graphics such as bar charts, line graphs, and pie charts play an increasingly important role in many documents found on the Web. These inherently visual constructs enable viewers to quickly and easily perform complex tasks such as comparing entities or identifying trends. Unfortunately, the growing use of visual information displays has disenfranchised individuals with sight impairments. Thus the challenge is to develop techniques for providing effective access so that all individuals can utilize these information resources in their education, work, decision-making, and leisure.

Many documents on the Web can be accessed by visually impaired users through screen reading software that reads the information on a computer screen using synthesized speech. If the developer of the web page has supplied alternative text (or “alt text”) for graphics in the HTML, most screen readers will read this text to the user. However, the vast majority of web pages are developed without accessibility in mind, and alt text is not supplied, thus making the content of the document’s graphics inaccessible to a visually impaired user. As one of our visually impaired users noted, “Of course, half the time they [web pages] don’t even label [the graphics] to say this is a chart, or, it just says ‘Graphic207’.”

2 Approach and Related Work

Our work is concerned with providing access to information graphics from popular media such as newspapers and magazines. Such information graphics generally have a message that the graphic designer intended to convey. For example, given the graphic in Figure 1, the graphic’s message is that there is an increase in Jan ’99 in the dollar value of 6-month growth in consumer revolving credit in contrast with the decreasing trend from July ’97 to July ’98. Rather than providing the visually impaired user with alternative access to what the graphic looks like or a listing of all of the data points contained in the graphic, our work takes a very different approach and attempts to provide users with the message and knowledge that one would gain from viewing the graphic. We eventually envision the system as an interactive system that will provide a richer textual summary, and that can respond to requests for further information about the graphic.

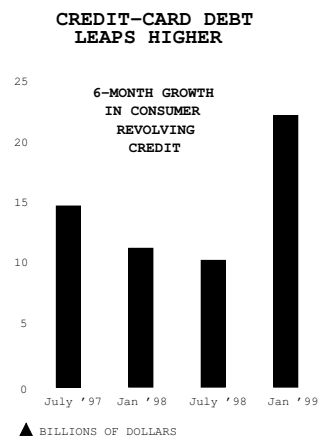


Figure 1. Graphic with a Contrast Pt with Trend Message¹

A number of projects have attempted to make graphs accessible to visually impaired viewers by reproducing the image in an alternative medium, such as sound [1], touch [10] or a combination of the two [11, 14]. Aside from the use of sound and touch, there is some research involving presenting graphics via text. For example, Ferres et al. [9], describe the inspectGraph system, which has the goal of providing accessibility for blind users to the graphs

¹Graphic from BusinessWeek, April 5, 1999 issue.

published in Statistics Canada’s “The Daily.” When the graph creator saves the graph as part of an Excel spreadsheet, a plug-in generates the required inspectGraph files. inspectGraph supports the communication of the contents of graphics in two ways: 1) it generates short, static textual descriptions of the content and appearance of the graphic that can be referred to by tags in HTML files, and 2) there is a navigation tool that allows users to query and navigate the content of the graph.

However, all of these approaches require the user to build a “mental map” of the diagram – a task that is very difficult for the congenitally blind because they have no personal knowledge regarding the appearance of information graphics [11]. In addition, many of the other systems require 1) special equipment or 2) preparation work (such as model creation) by a sighted individual. For example, inspectGraph [9] requires that support files be generated by a plug-in when the graph is created, and also requires the user to construct a mental map of the graph in order to infer any intended message. Consequently, existing systems have not been successful in solving the graph accessibility issue for visually impaired individuals. Thus it is imperative that novel approaches be investigated.

3 The Importance of Making Information Graphics Accessible

In order to assess the relative importance of being able to access the information contained within an information graphic, we conducted a corpus study [3] whose primary goal was to determine the extent to which the message conveyed by an information graphic in a multimodal document is also conveyed by the document’s text. We analyzed 100 randomly selected graphics from our collected corpus of information graphics, along with the articles in which they appeared. The selected articles were taken from magazines and local and national newspapers. In 39% of the instances, the text was judged to fully or mostly convey the message of the information graphic. However, in 26% of the instances, the text conveyed only a little of the graphic’s message. Most alarming was the observation that in 35% of the instances, the text failed to convey any of the message.

These findings are applicable when considering the broader issue of the accessibility for visually impaired users of all images on the Web. For example, the WebInSight [2] project is aimed at providing useful alternative text for images found on the web, such as those found in navigation bars, as form buttons, and to display textual and visual content. WebInSight attempts to extract any text found in the images and supply it to the user via the alternative text attributes of the img tag in the HTML. This method shows great promise for many critical navigational and functional images (such as those found in menus or on buttons), since the important text is often embedded in the image. However, in light of the corpus study described above, the text found within information graphics is un-

likely to consistently supply helpful information about the contents of the graphic. One could envision a system such as ours being used to generate the alternative text for information graphics within a framework such as WebInSight.

The remainder of the paper presents our implemented system for enabling visually impaired users to gain access to the information provided by simple bar charts that appear on a web page. The focus of this paper is on the interaction of the user with the system interface, the impact of the design decisions upon the user experience, and the user evaluations and feedback.

4 System Architecture

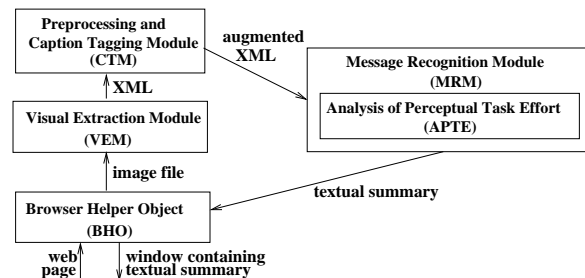


Figure 2. System Architecture

The architecture of our system is shown in Figure 2. The following subsections discuss the various components with particular emphasis on the browser extension which allows visually impaired users to access textual summaries of information graphics. Note that while the browser extension will work for any type of information graphic, the scope of the work currently implemented for the image processing and message inference components is limited to the processing of simple bar charts. By simple bar charts, we mean bar charts that display the values of a single independent attribute and the corresponding values for a single dependent attribute. However, we believe that our methodology is broadly applicable and extensible to other types of information graphics.

4.1 Browser Extension

We considered the following three goals while designing our system: 1) the system should be usable by as many visually impaired individuals as possible, 2) the system should not require the use of special equipment or assistance from a sighted helper, and 3) the interface presented to users should be easy to use and compatible with their navigation preferences. Each of these design goals influenced the form and functionality of the browser extension that serves as our user interface.

In order to achieve the first goal, reaching the largest potential audience of users, we have implemented our browser extension specifically for Internet Explorer, since it currently holds the majority web browser market share. We then tested it using Freedom Scientific’s JAWS, which

holds 65% of the screen reader market share [13]. However, the concepts applied here are extensible to other implementation platforms (see Section 4.1.3).

With respect to the second design goal, visually impaired computer users are already encumbered by having to buy relatively expensive screen reading or magnification software. Any solution which requires additional special equipment would likely limit potential users, as well as hinder the portability of the computer system it is attached to. Thus, being able to utilize our system with only a web browser and screen reading software is a strong advantage. The ability of our system to automatically infer the intended message of an information graphic ensures that intervention, in the form of assistance or preparation work, by a sighted individual (including the developer of the web page) is not necessary.

The final goal for our browser extension was that it should be as easy as possible for users to find and select an information graphic, signal for the system to infer the graphic's summary and convey it to the user, and then return to the original position in the web page. For this reason, our system is completely keystroke driven and utilizes keys that do not interfere with any of the current navigational settings for JAWS.

4.1.1 Interacting with the System

When navigating a web page, JAWS users have many options. When the web page is initially opened, JAWS begins reading the content of the web page, from top to bottom. The actual content that JAWS reads is highly configurable by the user, but typically includes any text on the page, the screen text pertaining to links and buttons, and the alternative text associated with graphics. Additionally, users could choose to press the "tab" key to traverse the content of the page, use quick navigation keys (such as G and Shift+G), or use the cursor (arrow) keys to control their navigation through the content of a web page. In order to avoid conflicts with the existing navigation commands in JAWS, we chose CONTROL+Z as the key combination for launching our system. If the user comes across a bar chart during their navigation of a web page, they can hit CONTROL+Z to receive a textual summary of the information conveyed by the bar chart. For example, if the user encountered the graphic shown in Figure 3, they could hit CONTROL+Z and a dialog box containing the summary of the graphic would appear. For this particular graphic, our system produces the summary "This bar chart titled 'The notebook spiral' shows that the dollar value of average laptop prices fell from 2000 to 2003 and then falls more slowly until 2007." By default, JAWS will read the contents of the dialog box aloud as soon as it is displayed.

This type of interaction requires a very tight coupling between our application and the web browser, because our application needs to be able to determine which graphic

THE NOTEBOOK SPIRAL

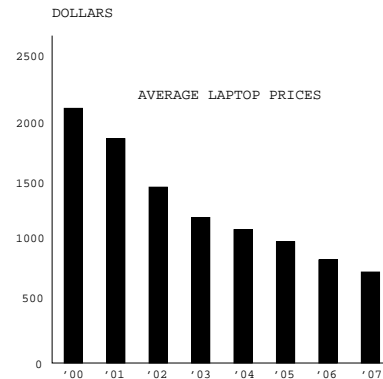


Figure 3. Graphic with a Changing Trend Message²

is currently in focus within the web browser and listen for user keystrokes. We achieved the proper level of integration by implementing our user interface as a Browser Helper Object for Internet Explorer.

Browser Helper Objects are special add-on components that enable the customization of Internet Explorer (version 4.0 or later). BHOs are tied to the main window of the browser and are created as soon as a new browser window is created. BHOs are implemented as in-process Component Object Model (COM) components, and they run in the same process space as the browser; this means that they can perform virtually any action on the available windows and modules. Our BHO hooks onto Internet Explorer's message loop and captures all of the keyboard events within the browser. Upon detecting the CONTROL+Z combination, if the BHO determines that the current focus in the Document object of the Internet Explorer instance appears to be a graphic containing a bar chart, our system attempts to infer the intended message of the bar chart.

4.1.2 Identifying Bar Charts

When a user navigates to a graphic, JAWS attempts to identify the graphic by using the "alt text" (if present) or the file name. However, these sources often do not contain any indication that the image represents a bar chart. Not surprisingly, some users have stated in their user evaluations that they often do not know when an image represents an information graphic. Therefore, as soon as the web page has loaded, our browser extension scans all the images on the page for images that appear to be bar charts. The scan uses simple tests, such as whether the graphic has fewer than 20 gray levels, and whether or not rectangles with aligned edges (as they would be aligned along an axis) are present in the image. If an image is considered to be likely be a bar chart, the browser extension appends the message, "This image appears to be a bar chart. Press Control+Z to determine the summary." to any existing alt text for the image. JAWS then reads this text to the user when the user navigates to the image.

²Graphic from BusinessWeek, September 5, 2005.

4.1.3 Extensibility of the Browser Extension

While the current version of the user interface has been designed specifically with JAWS and Internet Explorer in mind, we expect similar solutions to work for other applications. For example, extensions similar to BHOs can be developed for Mozilla's Firefox browser using the Cross Platform Component Object Model (XPCOM). Regarding the use of screen readers other than JAWS, our BHO in Internet Explorer will work with any screen reader; it is simply a matter of investigating how the focus of Internet Explorer and the screen reading software interact and of ensuring that the keystroke combination does not conflict with existing screen reader functionality. For visually impaired users who primarily use a screen magnifier (such as ZoomText), the text produced by our BHO can be handled in the same manner as text in any other application.

4.2 Processing the Image

After our system is launched by the user through the browser extension, the image is processed by the Visual Extraction Module. VEM is responsible for analyzing the graphic's image file and producing an XML representation containing information about the components of the information graphic including the graphic type (bar chart, pie chart, etc.) and the textual pieces of the graphic (such as its caption). For a bar chart, the representation includes the number of bars in the graph, the labels of the axes, and information for each bar such as the label, the height of the bar, the color of the bar, and so forth [5]. This module currently handles only electronic images produced with a given set of fonts and no overlapping characters. In addition, the VEM currently assumes standard placement of labels and axis headings. Work is underway to remove these restrictions. But even with these restrictions removed, the VEM can assume that it is dealing with a simple bar chart, and thus the problem of recognizing the entities is much more constrained than typical computer vision problems.

The XML representation is then passed to the Pre-processing and Caption Tagging Module (CTM). The pre-processing augments the XML with salience information such as a bar that is colored differently from other bars in the graphic or a bar that has an annotation when the other bars do not. The caption tagging extracts information from the caption (discussed later) and then passes the augmented XML representation to the message recognition module (MRM), which is described in the next section.

4.3 MRM: A Bayesian Inference System

We have developed a Bayesian inference system for recognizing the intended message of an information graphic. We hypothesize that this message can serve as the core content of an effective summary of the information graphic. This message inference methodology has several potential domain applications aside from accessibility for blind users,

such as generating searchable summaries of information graphics for digital libraries, and tutoring systems aimed at improving the design of information graphics. Note that this paper is centered on the issue of making bar charts accessible to blind users, and therefore this section is intended as a brief overview of our Bayesian inference system; further details on the network design, implementation and evaluation can be found in [8].

We view information graphics that appear in popular media as a form of language with a communicative intention. Therefore, we have extended plan inference techniques that have successfully been used in recognizing the intention of natural language utterances (for example [15, 4]) to the novel domain of information graphics. A critical component of any plan inference system is the evidence, or communicative signals, that are identified and utilized by the system to infer the communicative intention of an agent. In this case, the MRM is attempting to infer the message that the graphic designer intended a viewer of the graphic to recognize. We have identified three categories of communicative signals that appear in simple bar charts.

Our first communicative signal is the salience of entities in the graphic. The graphic designer can employ a number of strategies for making an entity or entities in a bar chart salient to the viewer. For example, a bar could be shaded or colored differently than the other bars (as is the case for the bar labelled CBS in Figure 4), or could be specially annotated. Mentioning a bar's label in the caption of the graphic also makes the represented entity salient, since this draws attention to the bar.

A second communicative signal is the relative effort required for different perceptual and cognitive tasks. Here we are extending a hypothesis of the AutoBrief group [12]. The AutoBrief project was concerned with generating information graphics, and they hypothesized that a graphic designer chooses a design that best facilitates the perceptual and cognitive tasks that a viewer will need to perform on the graphic. We posit that if a graphic designer went to the effort of making a particular perceptual task easy to perform on a given graphic, then that perceptual task is likely to be part of the set of tasks that the viewer was intended to perform in deciphering the graphic's message.

A third communicative signal, and the only one external to the information graphic itself, is the presence of certain verbs and adjectives in a caption. In [7] we present a corpus study showing that (1) captions are often very general or uninformative, and (2) even when captions convey something about the graphic's intended message, the caption is often ill-formed or requires extensive analogical reasoning. Therefore, rather than attempting to actually *understand* the caption, we utilize shallow processing of the caption in order to extract communicative signals.

The communicative signals that we extract from an information graphic are then utilized by our system to reason about the intended message of the graphic. For each new information graphic, we dynamically construct a Bayesian network. The top level of the network captures

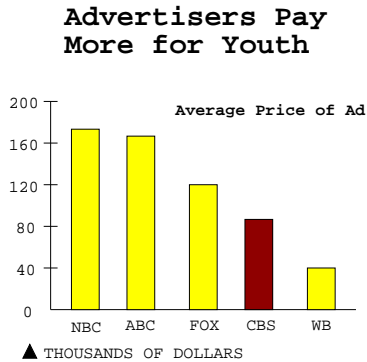


Figure 4. Graphic with a Get-Rank Message⁴

the various categories of messages that can be conveyed by a bar chart, such as conveying a change in trend (Change-Trend), conveying the rank of an entity in a bar chart (Get-Rank), comparing two entities (Relative-Difference), and so forth.³ The communicative signals extracted from a graphic are represented in the Bayesian network as evidence nodes. Once the network is constructed, the probabilities propagate through the network to hypothesize the intended message of the graphic. For the graphic in Figure 4, the system infers that the graphic is conveying the rank of CBS and produces the summary “*This bar chart titled ‘Advertisers pay more for youth’ shows that CBS has the second lowest rank in terms of the dollar value of average price of Ad compared with NBC, ABC, FOX, and WB.*”

4.4 Generating the Summary

Once the intended message has been inferred by our Bayesian inference system, it is used as the core content of a textual summary of the graphic. Generating a coherent natural language summary is a non-trivial issue; for example, one of the most challenging aspects of generating coherent natural language has been determining the full label for the measurement (or value) axis. Further details on realizing the inferred intention in natural language can be found in [6].

5 User Evaluations

Of all of the data that could be collected regarding a system’s performance, user evaluations are the most critical, particularly when dealing with users with disabilities. In our user evaluations, participants were first given a guided introduction to our system. Participants then explored various web pages containing bar charts, using our system to access summaries of the bar charts. The evaluation concluded with a taped interview. Two areas were assessed: the effectiveness of our overall methodology in enabling

³Other messages include conveying the relative rank of all entities, conveying trends (increasing, decreasing, or stable), contrasting a data point with a trend, and conveying which entity has the maximum or minimum value in a graph.

⁴Graphic from BusinessWeek, April 5, 1999.

visually impaired users to access the content of bar charts, and the ease of use of the browser interface.

Ten visually impaired users participated in our study. The participants had varying degrees of computer experience, though all were JAWS users. The participants also possessed varying degrees of vision, but none could view the graphics without substantial magnification (which was not used during the experiments). There was also a mix of congenitally and non-congenitally blind users.

Three of the evaluation questions were numerical in nature. The responses from each of the users is shown in Table 1. The questions were:

- On a scale of 1–10 with 1 being not useful at all and 10 being extremely useful, how useful would software be that can provide the message and content of a graph?
- On a scale of 1–10 with 1 being extremely difficult and 10 being extremely easy, how easy was it to select the graph within the web page?
- In order to start our system, you had to press Control+Z. On a scale from 1–10 with 1 being extremely difficult and 10 being extremely easy, how easy was this to do?

User	Impairment	Usefulness	Ease of Use	Ease of Keystroke
1	Congenital	6	8	No answer
2	Congenital	10	10	10
3	Congenital	10	10	10
4	Congenital	10	10	10
5	Legally	8	9	9
6	Congenital	5	10	10
7	Noncongenital	9	5	10
8	Noncongenital	9	9	9
9	Noncongenital	10	10	10
10	Congenital	5	10	10
Avg		8.2	9.1	9.8
Stdev		2.1	1.6	0.4

Table 1. Quantitative User Evaluation Results

It is clear from the results of the user evaluations that the participants were pleased with the overall methodology and ease of use of the system. The response from the participants was overwhelmingly positive, and the participants were generally very excited about the possibility of being able to access the content of information graphics. One participant remarked, “... sometimes the information is what you’re trying to get. If you go over that graphic and then you run through this whole article or document and you’re not getting what you want, it might be in that little bit. You know, which has happened in other cases.”

One of our open-ended questions was whether or not the participants would like to be able to obtain further information from the graphics, and if so, what information they would like to be able to get. A striking, though in retrospect not surprising, difference emerged between the congenitally and non-congenitally blind users as to whether they could identify what other types of information they might like to have. This inherently makes sense, since the congenitally blind users have never had the opportunity to

view or use bar charts. As one participant noted, "... I don't know enough about graphs to answer that question. Having never been able to see." We posit that this distinction lends credence to the hypothesis that it would be difficult for congenitally blind users to form a mental map of a bar chart, based on an alternative description of its appearance, in order to extract the message of the graphic. Although many congenitally blind participants felt that the system would be quite useful, the three who rated the system's potential usefulness below a 7 were congenitally blind and had never viewed a bar chart. We surmise that these lower ratings reflect their lack of familiarity with information graphics and the valuable information that they provide.

Another very interesting result of the evaluations was that when asked "Did you have confidence that the output from our system helped you to understand what the graph was conveying?" all of the participants responded positively. While encouraging, this result also underscores the responsibility of a system to avoid misleading its users. It should be noted that even if our system incorrectly infers the graphic designer's intended message for a bar chart, the inferred message still reflects information that is present in the graphic.

6 Conclusion

Information graphics are an important part of many documents available on the world-wide web, yet they are largely inaccessible to visually impaired users. This is clearly unacceptable. As one of our participants stated, "I think that we deserve to have as much information as everyone else does." This paper has presented a novel implemented interface that enables visually impaired users to gain access to the information provided by simple bar charts that appear on a web page. Our approach of presenting the message conveyed by the information graphic does not require specialized hardware, preparation work by a sighted individual, or for the user to construct a mental map of the graphic. Our user evaluations indicate that our approach shows great promise for addressing the challenge of providing universal access to information graphics.

References

- [1] J. L. Alty and D. Rigas. Communicating graphical information to blind users using music: The role of context. In *Proc. of CHI-98, Human Factors in Computer Systems*, p. 574–581, Los Angeles, 1998. ACM Press.
- [2] J. P. Bigam, R. S. Kaminsky, and R. E. Ladner. WebInSight: Making web images accessible. In *Proc. of ASSETS 2006*, p. 181–188, October 2006.
- [3] S. Carberry, S. Elzer, and S. Demir. Information graphics: An untapped resource for digital libraries. In *Proc. of SIGIR 2006*, Seattle, WA, 2006.
- [4] E. Charniak and R. P. Goldman. A Bayesian model of plan recognition. *Artificial Intelligence*, 64(1):53–79, November 1993.
- [5] D. Chester and S. Elzer. Getting computers to see information graphics so users do not have to. In *Proc. of the 15th Int'l Symposium on Methodologies for Intelligent Systems (ISMIS), LNAI 3488*, pages 660–668. Springer-Verlag, 2005.
- [6] S. Demir, S. Carberry, and S. Elzer. Effectively realizing the inferred message of an information graphic. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP)*, pages 150–156, 2007.
- [7] S. Elzer, S. Carberry, D. Chester, S. Demir, N. Green, I. Zukerman, and K. Trnka. Exploring and exploiting the limited utility of captions in recognizing intention in information graphics. In *Proc. of the 43rd Annual Meeting of the Assoc. for Computational Linguistics*, pages 223–230, June 2005.
- [8] S. Elzer, S. Carberry, I. Zukerman, D. Chester, N. Green, and S. Demir. A probabilistic framework for recognizing intention in information graphics. In *Proc. of the 19th Int'l Joint Conf. on Artificial Intelligence*, pages 1042–1047, July 2005.
- [9] L. Ferres, P. Verkhogliad, G. Lindgaard, L. Boucher, A. Chretien, and M. Lachance. Improving accessibility to statistical graphs: the InspectGraph system. In *Proc. of ASSETS 2007*, Tempe, AZ, October 2007.
- [10] S. Ina. Computer graphics for the blind. *ACM SIG-CAPH Computers and the Physically Handicapped*, 55:16–23, 1996.
- [11] A. R. Kennel. Audiograf: A diagram-reader for the blind. In *2nd Annual ACM Conf. on Assistive Technologies*, p. 51–56, 1996.
- [12] S. Kerpedjiev and S. Roth. Mapping communicative goals into conceptual tasks to generate graphics in discourse. In *Proc. of the Int'l Conf. on Intelligent User Interfaces*, p. 60–67, 2000.
- [13] C. Law, J. Jacko, and P. Edwards. Programmer-focused website accessibility evaluations. In *Assets '05: Proc. of ASSETS 2005*, p. 20–27, New York, NY, USA, 2005.
- [14] D. K. McGookin and S. A. Brewster. Soundbar: exploiting multiple views in multimodal graph browsing. In *NordiCHI '06: Proc. of the 4th Nordic Conf. on Human-computer Interaction*, p. 145–154, New York, NY, USA, 2006. ACM Press.
- [15] R. Perrault and J. Allen. A Plan-Based Analysis of Indirect Speech Acts. *American Journal of Computational Linguistics*, 6(3-4):167–182, 1980.