July 22, 2007

RT0715
Human-Computer Interaction 10 pages

Research Report

Technology Advances and Standardization toward Accessible Business Graphics

Hironobu Takagi, Tatsuya Ishihara

IBM Research, Tokyo Research Laboratory
IBM Japan, Ltd.
1623-14 Shimotsuruma, Yamato
Kanagawa 242-8502, Japan

Technology Advances and Standardization toward Accessible Business Graphics

Hironobu Takagi[1], Tatsuya Ishihara[1]

[1] Tokyo Research Laboratory, IBM Japan, 1623-14, Shimo-tsuruma, Yamato, Kanagawa, Japan, 242-8502
{takagih, tisihara}@jp.ibm.com

Abstract. Various types of graphics are exchanged in our daily business and education processes. In spite of the importance of business graphics, they are not accessible for visually impaired people, especially for the blind, and this impacts their productivity at work. Current serialization-based screen reading techniques do not provide sufficient functionality for accessing graphics. In this paper, we report the results of our survey to identify next generation accessibility features in future graphics standards, especially for the OpenDocument Format (ODF). We will first compare accessibility functions in various types of existing standards. Then we will report our survey results for three related areas, a survey of existing business graphics in presentation documents to unveil the complexity of practical graphics, a survey of research on various types of graphical taxonomies, and a survey of interface technologies for representing graphics non-visually, such as screen reading and pictorial Braille. Finally, we will propose three practical enhancements for standard graphic formats based on the survey results.

Keywords: Accessibility, graphics, presentation, visually impaired, standards, OpenDocument Format.

1 Introduction

Various types of graphics are exchanged in our daily business and education processes, such as organizational charts, process diagrams, timeline charts, software architecture charts, mathematical and statistical charts, and so on. These business graphics convey vital information and knowledge about business projects, and they should be shared among people who work together on or in the projects.

In spite of the importance of business graphics, they are not accessible for visually impaired people, and thus impact on their work productivity. This barrier is becoming severe as business visualization extends deeper into business processes. Currently, no standardized conversion method from visual graphics into non-visual media is known, since the bandwidth of vision is incomparably wider than any non-visual senses, such as the senses of hearing or touch.

Currently, screen readers and voice browsers are the major methods to access arbitrary documents, but they do not have sufficient functionality for accessing graphics. Screen reading is a method to read aloud (or display one line of Braille on a refreshable Braille device) focused content fragments according to users' keyboard operations. Blind users are required to recognize the underlying visual structure by exploring among the graphical objects by using a keyboard. In other words, the graphics are serialized into a sequence of text fragments, and users need to imagine the corresponding visual structures through a keyboard-based exploration process instead of eye-movement-based visual exploration.

As members of the OASIS OpenDocument Format (ODF) accessibility sub-committee, we have been working to improve the accessibility of standardized office document specifications including those for presentation documents. The presentation document specification in the latest ODF (version 1.1) has basic accessibility functions, such as alternative text for a group made of graphical primitives, graphics-caption relationship assignment, and reordering functions for tab navigation. These functions are comparable to any other vector graphics format, but it is clear that these functions are insufficient to make most business graphics accessible.

In this paper, we report on our survey to help design next generation accessibility features in the future versions of the ODF presentation documents specifications. First we will compare accessibility functions in various types of existing standards. Then we will report our survey results for three related areas, research on graphical taxonomies, existing business graphics, and interface technologies. Finally, practical future improvements to the standards will be discussed.

2 Existing Accessibility Standards for Business Graphics

Table 1 shows a comparison of existing standards for documents and graphics, and their considerations regarding accessibility, especially for the blind. We investigated five content standards, W3C HTML 4.1, W3C/OASIS WebCGM 2.0, W3C SVG (Scalable Vector Graphics) 1.2, ECMA Office Open XML 1.0, and OASIS OpenDocument Format (ODF) 1.1. A draft metadata standard, W3C ARIA, is also shown in the table for comparison. It is designed to complement W3C content types (HTML, WebCGM, and SVG) in terms of dynamic Web content accessibility.

We classified accessibility functions into seven categories, five categories for text descriptions, and one category each for object groupings and navigation orders. In other words, the union of the current accessibility functions in the various standards only covers these concepts for accessible graphics. The goal of this paper is to identify any missing concepts and to investigate future possibilities for enhancements.

"Text descriptions" are the basic method to make graphics accessible for screen readers by adding text alternatives. Most of content types have functions to add at least two levels of invisible text descriptions (short and long) to each graphical object. “Speaker notes” are not officially defined for accessibility purposes, but are widely used to describe presentation slides. Only ODF has a function to associate a visible text with a graphical object as a description (“Visible label”). HTML, SVG, and WebCGM will have association functions, since the ARIA draft has two types of metadata for the association (describedby and labelledby).

“Object groupings” allow content authors to mark a set of graphic elements as a semantically indivisible object and to add text descriptions to the group of objects. This function can also be used to create hierarchical structures by adding nested groupings. All vector graphics types have grouping functions, and ARIA defines several types of object groupings.

Since screen-reading is the basic method for accessing graphics, each content type has functions to compensate for the order of screen-reading (“Navigation orders”). In order to mark the orderings, we found three types of methods for ordering objects. ARIA’s “tabindex” is an attribute to assign a priority value to

Table 1. Comparison of existing graphic standards in terms of accessibility for the blind.

			W3C HTML4.1	W3C/OASIS WebCGM 2.0	W3C SVG 1.1	ECMA Office Open XML 1.0	OASIS ODF1.1	W3C-ARIA (draft)
Graphics type			images	images and vector graphics	images and vector graphics	images and vector graphics	images and vector graphics	
Attributes / Elements for Accessibility	Text descriptions	Short descriptions	ALT	content	title	alt(VML), descr(DrawingML)	svg:title	
		Long descriptions	LONGDESC	layerdesc	desc		svg:desc	
		Regional descriptions	MAP & ALT				draw:image-map & svg:desc	
		Speaker notes	NA	NA	NA	p:notes	presentation:notes	
		Vsible labels					draw:caption-id	describedby, labelledby
	Objects groupings		NA	grnode	g	p:grpSp	draw:g	group, region, section, list, etc.
	Navigation orders		NA				draw:nav-order	tabindex, posinset, flowto, etc.

each object. It is easy to adjust the reading orders among a small number of objects, but it tends to be difficult to control large numbers of objects. ARIA's "flowto" is a method to point to a next object to set a relevance-based reading order, but improper assignments can cause inappropriate skipping of content. ODF's "draw:nav-order" is a method to describe a sequence of object IDs in a property list. This method is simple and easy to manage, but no interface to change this attribute has yet been implemented in any office document editor. Therefore, though the idea seems promising, it is still too early to assess its utility.

3 Survey

This survey aims at identifying important missing functions in the current standards. Three types of different surveys were conducted to cover a wide range of related knowledge, and to anticipate future accessibility environments.

3.1 Survey of Existing Business Graphics

The first approach was an analysis of existing business graphics. We collected more than 4,000 pages of presentations from the Internet and also various types of graphics used in our company, and then selected about 60 pages as typical examples. Fig. 2 shows an overview of these charts. The horizontal axis shows the complexity of the graphics and the vertical axis shows the formality. We loosely categorized these graphics into four categories. "Simple graphics" are relatively simple, so we think that these graphics can be translated into accessible formats in various ways, such as pictorial Braille or alternative voice access interfaces (see Section 3.3).

"Formal diagrams" are formally defined diagrams, which have a well-defined data-model behind the visual structure, such as tree structures, Venn diagrams, cause-and-effect diagrams, and UML (Unified Modeling Language) diagrams. If a non-visual conversion were defined for each data model and blind users learned the appropriate access method, then these diagrams should be accessible. In order to implement this approach, the graphics should have functions to preserve the data-model for each part of the figure.

"Table structures" are graphics that can be regarded as tables. If the visual information of a graphic can be represented as a table, then blind users can access them by using table navigation functions in screen readers. In this case, standards would need to have functions to add tabular metadata to the graphical objects.

"Composite diagrams and graphics" have multiple types of diagrams in one figure, annotated formal graphics, or visually represent multiple aspects (dimensions) of some information in one figure. In the case of Test 46 (Fig. 2), the basic diagram structure shows a history of topic changes and each vertical box indicates a year. These two aspects are well represented in the diagram, but it is

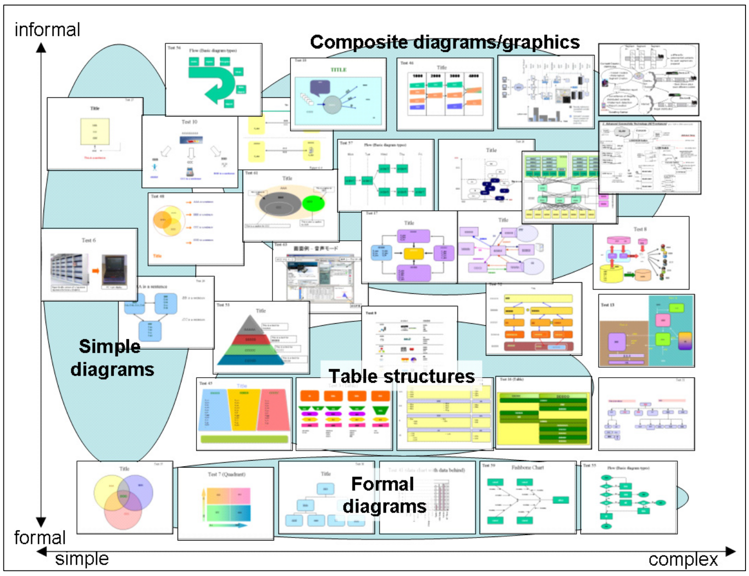

Fig. 1. Variety of collected existing graphics in presentation documents.

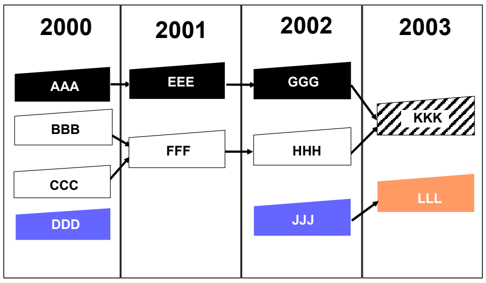

Fig. 2. Example of a graphic with multiple aspects. The color scheme indicates categories of topics (nodes), the horizontal positions show a timeline, and the arrows indicate relationships between the topics.

difficult to represent these aspects non-visually. At this time, general methods to make them accessible are not well known. In fact, more than 50% of the investigated graphics were classified into this category.

3.2 Survey for Graphical Taxonomies

That survey of existing graphics showed us the complex nature of practical graphics. It raised questions about the possibilities of making them accessible, since they contained various graphical objects with various visual properties and they were aligned to express specific knowledge and information. In this survey, we will try to figure out possibilities to describe complex graphics by referring to research on graphical taxonomies.

Lohse et al. [1] classified various types of graphics into eleven categories based on the results of user studies, such as “structure diagrams”, “cartograms”, “maps”, “graphic tables”, “process diagrams”, “icons”, “time charts”, “network charts”, “pictures”, “tables”, and “graphs”. Their coverage of graphics types is larger than our survey as described in the previous section, since they collected a variety of graphics from popular reference books on graphics. Their classification and user studies show us the breadth of the problem space here.

Blackwell and Engelhardt [2] conducted the most comprehensive survey for diagram research, and it resulted in a set of meta-taxonomies. This meta-taxonomy yields a useful framework for the necessary metadata in presentation documents. Among their nine taxonomic aspects, at least six representation-related aspects have a close relationship with metadata for accessibility, from “graphic primitive elements and the graphic properties” to “represented information”. Between “graphic primitive” and “represented information”, two aspects for symbols (“conventional elements” and “pictorial abstraction”) are defined, and “graphic structure” is an aspect for the two-dimensional structure of graphics, such as linear sequence, two-axis-chart, table, or tree structure. The “mode of correspondence” is a unique aspect, which describes the relationship between a representation and its meaning, such as literal vs. metaphorical, direct vs. indirect, and iconic vs. symbolic.

It is clear that the “represented information” is the most important information to be presented non-visually, however no standards support bundling of the represented information in graphics. The idea of “data-model attachment” is an important idea (See Section 4). It is not clear whether “mode of correspondence” can help non-visual access or not, so we need to investigate this aspect further. The “graphic structure” should also be one of the focuses of future standardization, since it helps translate into non-visual formats. The aspects for symbols are handled well in current standards as various levels of text description. The “graphic primitives” are less important, since they are already preserved in the file formats as objects and properties.

Kong et al. [3] conducted a survey of low-level taxonomies for graphic structures. Since they aimed at defining a formal graphical grammar for their visual language system, the taxonomy only includes metadata suitable for automatic detection, such as topology (such as “touch” and “overlap”), direction (an 8-point compass), distance (close, far) and alignment (such as “gap”, “middle”, “left” and “right”). The SVG RDF proposal [4] also includes such low-level metadata, such as topology (such as “Contains”, “isPartOf” and “InFrontOf”), direction (InPosition, InDirection), and alignment (AtLeft, AtRight, OnTop, Under). The necessity of this low-level metadata for accessibility purposes needs to be considered, since it would impose heavy work requirements on the creators of graphics, and the benefits for accessibility are not clear.

3.3 Survey for Interface Technologies

The third survey is of the technologies of media conversion for sensory substitution. This survey lists conversion methods that are accessed by two different human sensory systems, screen reading for the sense of hearing and pictorial Braille for the sense of touch.

3.3.1 Screen Reading

The one-dimensional speech interface is widely used and speech can even describe structural metadata information with interactions. For example, table navigation allows users to recognize two-dimensional table structures by using various commands, such as directional movement (N, E, S, W), table header reading, and consistent location reading when a user changes column or row. However, it is clear that the accessible structures are limited to relatively simple structures, and insufficient for the complexity of general business graphics.

As a trial to improve screen reader navigation, we developed a presentation access tool optimized for screen reader users, which has automatic transformation functions. We call this interface “DocExplorer” [5]. We implemented automatic detection functions for basic metadata taxonomic items, such as nodes, arrows/arcs, and groups, by analyzing primitive graphical elements. Then the tool creates a tree-view structure, which is a popular GUI control to represent hierarchical data (and familiar to many blind users).

Through the development of the automatic detection algorithm, we found various problems in improving the detection accuracy, even for such basic metadata. The accuracy could be greatly improved (or degraded) by small graphical changes. For example, automatic detection of topological metadata for grouping, such as “in”, “crossing”, or “touching”, was affected by the precise positioning of the shapes. These results indicated a need to manually create the low-level metadata (See Section 3.2), but the workload of authoring needs to be taken into account.

We also evaluated our tool with several blind employees in our company. Considering their prior experiences, they evaluated the tool as the best text conversion tool for presentation documents. However, the reliability of the depicted structures was raised as a problem. We believe this result indicates limitations of the current technology, and also the necessity for new meta-data standards.

3.3.2 Pictorial Braille

The two dimensional tactile interface is also a common method used for reading graphical content. Today, Braille printers are commercially available, and they are in widespread use in many facilities. However, the resolution of Braille pictures is quite low compared to the original graphic pictures, because of the limitation of the human sense of touch. While the resolution of a typical laser printer is 600 dpi, the resolution of a Braille picture is about 20 dpi. For this reason, a manual conversion process is required to create Braille pictures. It takes a lot of empirical knowledge to convert graphics to Braille pictures [6].

Currently great efforts are being made to improve the tactile interfaces, and remarkable progress is being made. One newly created interface is a touch panel which implements a coordinated presentation of a Braille picture with a computer. As the user touches a region of the Braille picture, a text label attached to it is read out by a computer. This interface is already commercially available, for example in IVEO® [7] and the Talking Tactile Tablet [8]. These are useful for understanding graphical content, including graphs, maps, and so on. Because there is a lot of graphical content in teaching materials and business presentation documents, these products are expected to be popular in schools and in business environments.

Another is the refreshable dot matrix display. This interface has been studied over the years [9]. Recently the price of this device is decreasing, making it commercially available. For example, the price of the DotView produced by KGS is about $10,000 (US) for a display size of 32 x 48 dots. It is used mainly for educational purposes, because the resolution is lower than Braille pictures.

No matter which interface is used, images rendered for the blind have much lower resolution, because the resolution of tactile perception is limited (about 1 dot/mm) [10]. That is why abstraction and simplification are necessary for creating useful Braille images [6]. Many activities are going on to make automatic conversion practical [11][12]. In order to help these manual and automatic conversions, metadata supporting the simplification should be considered in the standardization processes.

4 Toward Accessible Business Graphic Standards

Through these surveys, we can see the limitations of current technologies and standards. The path to accessible business graphics environment appears long and difficult and various activities are required. However, we found some hints about how to move toward the goal. In this section, we would like to propose three realistic improvements as the conclusion of this survey.

Mechanisms for attaching data models

The survey of existing graphics showed what kinds of graphics are commonly used, and the “formal diagrams” are among them (Section 3.1). The use of these formal diagrams is increasing in enterprises, such as UML diagrams [13] and Mindmaps. These well defined diagrams have clear data-model representations in XML or other formats. Since data-models already exist (which is why the graphics exist), no additional workload is required to attach data-models to these graphics in presentation documents. These data-models are corresponds to "represented information" in Blackwell and Engelhardts's meta-taxonomy (See Section 3.2). At this time, non-visual representations of these data-models are not well established, but the transformations for simplified user interfaces and pictorial Braille transformations would benefit from standardization (See Section 3.3).

Generalization of "connectors"

One of the most basic graphical structures in existing graphics is a directed connection between two shapes or images, which is represented using various types of arrows. Both of the presentation formats introduced in Section 2 (Office Open XML and ODF) have functions to draw directional arcs between two graphical objects. However, connectors are used only for drawing lines in the current standards. If connectors can be assigned to any type of arrow shape or image, then the connectors would become general metadata for visual connection structures. At this time, connectors are not utilized for non-visual access. It is clear that non-visual translations both for simplified user interfaces and for pictorial Braille can be improved by connector information (See Section 3.3).

This mechanism can improve the general usability for authors, since arrow graphics could follow shape relocations automatically. We believe that this is a kind of kill-two-birds-with-one-stone solution.

Preservation of alignment information

In Section 3.2, we introduced considerations of low-level metadata for accessibility. We need to take into account both the benefits of this metadata and the workload for authoring. The benefits are not clear, but non-visual translations, especially for pictorial Braille may benefit, since these techniques require drastic simplifications of the graphics based on the visual structures (See Section 3.3). Therefore, we would like to propose mechanisms to preserve alignments in documents recorded as metadata. Graphic authors naturally use alignment functions to layout graphic objects. Histories of these alignment operation could be used as hints for the automatic translation.

5 Conclusion

This paper proposed practical enhancements for future business graphics standards through a set of comprehensive surveys. First, we gave an overview of current graphics standards in terms of accessibility for the blind. Even for the best formats, accessibility considerations are limited by the functions of existing screen readers. Then, we first reported our survey results for existing business graphics. We collected a thousand pages of presentation images and analyzed them. We were able to organize them into four categories, but it was not clear how to make accessible the graphics in the largest category, the "composite graphics". The second survey was of research on graphical taxonomies. We found well defined frameworks for graphic accessibility. Two categories of technologies were reviewed in the third survey of interface technologies. The first approach is an expansion of current screen reading techniques, but it converts visual structures into suitable non-visual structures for screen readers. The other approach is pictorial Braille, such as paper Braille or refreshable dot matrix displays. These devices are becoming inexpensive and offer improved resolutions, but still cannot be compared to visual resolutions. That is why automatic transformation is a new and crucial technology to make them practical. Finally, three enhancements to standards were proposed. Data-model attachment will allow users to explorer the "represented information" underlying the graphics. The "generalization of connectors" will benefit both sighted graphic authors and blind end users. The "preservation of alignment information" is a low-hanging-fruit approach to add richer metadata to graphics to help with automatic transformations. We hope these proposals will be polished in the standardization process and will contribute to improve the accessibility of business graphics.

References

1. Lohse, G. L., Biolsi, K., Walker, N., and Rueter, H. H.: A classification of visual representations, Communications of the ACM, v.37 n.12, pp.36–49, Dec. 1994
2. Blackwell, A. F. and Engelhardt, Y.: A Meta-Taxonomy for Diagram Research, in Diagrammatic Representation and Reasoning, Anderson, M., Meyer, B., and Olivier, P., Editors. 2002, Springer, pp. 47–64
3. Kong, J., Zhang, K., and Zeng, X.: Spatial graph grammars for graphical user interfaces. *ACM Trans. Comput.-Hum. Interact.* 13, 2, Jun. 2006, pp. 268–307
4. Herman, I., and Dardailler, D.: SVG Linearization and Accessibility, Computer Graphics Forum, 21(4), 2002.
5. Ishihara, T., Takagi, H., Itoh, T., and Asakawa, C.: Analyzing visual layout for a non-visual presentation-document interface, *ACM SIGACCESS 2006*, pp. 165–172
6. Schuffelen, M.: On editing graphics for the blind, The Hague, Holland: Netherlands Library for Audio Books and Braille, 2002, http://www.stanford.edu/~sipma/grbl0.html
7. IVEO®, http://www.viewplus.com/products/touch-audio-learning/IVEO/
8. Talking Tactile Tablet, http://www.touchgraphics.com/ttt.htm
9. Kawai, Y. and Tomita, F.: Interactive tactile display system: a support system for the visually disabled to recognize 3D objects, *ACM SIGACCESS 1996*, pp. 45–50
10. Way, T. P., and Barner, K. E.: Automatic visual to tactile translation, part I: Human factors, access methods and image manipulation. *IEEE Trans. on Rehabilitation Engineering*, 5:81–94, 1997
11. Rotard, M., Knödler, S., and Ertl, T.: A tactile web browser for the visually disabled, *ACM HYPERTEXT 2005*, pp. 15–22
12. Ladner, R. E., Ivory, M. Y., Rao, R., Burgstahler, S., Comden, D., Hahn, S., Renzelmann, M., Krisnandi, S., Ramasamy, M., Slabosky, B., Martin, A., Lacenski, A., Olsen, S., and Groce, D.: Automating tactile graphics translation, *ACM SIGACCESS 2005*, pp. 150–157
13. Unified Modeling Language (UML), Version 2.0, http://www.omg.org/technology/documents/formal/uml.htm