Towards a Science of Document Intent

Steven J. Harrington
Xerox Digital Imaging Technology Center

Fernando Naveda
Rhys Price Jones
Rochester Institute of Technology

Abstract

The intentions of the document creator effect the decisions that go into the construction and presentation of the document. Preserving intentions across different display devices may require different decisions to be made. Intent information would be particularly useful to web documents where the creator does not control the choice of display device. An approach to quantifying intent based on statistical analysis of the value properties associated with decisions is described.

1. Introduction

When documents are constructed many decisions are made. The decisions occur at many levels and stages in the construction, from deciding what is to be said, whether to use text or pictures, what language to use, whether to use color, what font to use, how to layout the document elements, down to what halftone to employ and how out-of-gamut colors are handled. Some of the decisions are made by the author, while others may be built into the tools used to produce the document. The reasons behind the decisions are rarely captured explicitly. However, different intentions can result in different decisions, that in turn yield documents with very different appearance. For example, a section of my office phone list looks like figure 1, while a section of similar content from the phone book looks like figure 2.

The very different appearance is an indication of different intent. For the office phone list legibility was more important than economy, while for the phone book, reducing cost was a strong concern.

Electronic documents, connectivity and the Web have raised the need for understanding and capture of intent information. In old work processes, where the author had some control of the document decisions right up through its printing, it was sufficient for the intents to remain inside of the author's head. But when the document is distributed in electronic form, many decisions with respect to its presentation and use remain to be made, and the author is not around to influence them. For example, decisions about layout, font size, color and such may be influenced by whether the document is to be presented on a CRT, printed, displayed on a PDA or read to a cell phone. The author's intentions should be preserved across presentation media and devices, even though many of the actual decisions should change.

As an example of the effect of output device, consider the copyright or legal notices such as found for computer application software, as shown in Figure 3. In a hardcopy instruction manual, a legal notice may be located at the end of the document where it is easy to ignore. The font size may go from 10 or 12 point for the instructions to 6 or 8 point for the notice, and the line spacing may be decreased as well making it difficult to read. Sections may be printed in all caps, which gives it a feeling of importance while making it even harder to read.

Michael Brown	8*555-1213	128/283A
Jay Ciemore	8*555-3493	128/114
Sally Jones	8*555-3112	128/212E

Figure 1: Office phone list

BROWN Adam 57 Main St 14567….216-3334
John 234 Elm St 14567…….216-2274
Kevin & Sue
3231 Imperial Dr 14568…..…..322-2339
Lauren 21 Rockway Rd 14567…..216-9975
J 1343 Oak St 14567…..……216-3574
Peter S
3231 Wilson Dr 14568.……322-2339
BUDA Jas 57 Main St 14567…..216-3334
Wm 21 Rockway Rd 14567…..216-9975

Figure 2: City directory

6. Click **Restart Windows**

That's it! You have successfully installed the device

Copyright Information
Software release 3.4.5 for Windows. Copyright © 1998 SHWARE Inc. All rights reserved.
Reproduction, adaptation, or translation without prior written permission is prohibited except as allowed under the copyright laws.

Limited Product Warranty
IN NOEVENT WILL SHWARE BE LIABLE FOR DIRECT, INDIRECT SPECIAL, INCIUDENTAL OR CONSEQUENTIAL DAMAGES RESULTING FROM ANY DEFFECT IN THE PRODUCT OR FROM ITS USE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGES

Figure 3: Printed legal notice

Similar notices can be found displayed on workstations for downloaded applications. The output device does not support such tiny fonts, so this option is not available, but typing in all caps is still used. In addition, a small dialog window that cannot be resized is employed to display only a few lines of the text at a time, making reading uncomfortable. The dialogue box also has a big "Yes" button that will make it all go away.

This example shows that it is not always the author's intention to make the document comfortable to view, easy to read or convenient to use. Second, it shows that different output media require different expressions of intent. And third, it would be difficult to carry out an automatic transformation of the document for a target device without some understanding of the intent. (A dialogue box may indicate the need for action, but it the contained message one of urgent importance, or is it something that should be ignored).

The mechanism for presenting a phone list in a web browser should be different from how it is presented in a printed document. An approach currently being taken to deal with this problem is to separate of content from style and to employ style sheets. One can create style sheets for each class of output device with the idea that the style sheet will present the content in an acceptable way for that device. Note, however, that this is insufficient. Both phone list examples come from printers, but look very different and would not both be generated by a single style sheet for the printer. Style sheets can also be generated by the author, but this is awkward, because it requires the author to anticipate all the output devices on which the document might be presented. Since the presentation device and the author are be separated by time and space it is hard to generate a style sheet that comprehends both the device characteristics and the author's intentions. What is needed is a mechanism to capture the author's intentions as metadata in the document and have them available for use in the document's presentation. This is needed for electronic documents and web documents in particular since these are natural for distribution and display on a variety of devices.

We call the captured intentions the *intents*, and note that different portions of a document can have different intents. The intents associated with the cover, front matter, contents and body of a book are probably all different. Similarly, the intents associated with various frames of a web page may be different. Web document formats have mechanisms that can be used for associating metadata with document components. The problem lies in deciding what the intent metadata should be and how to extract it.

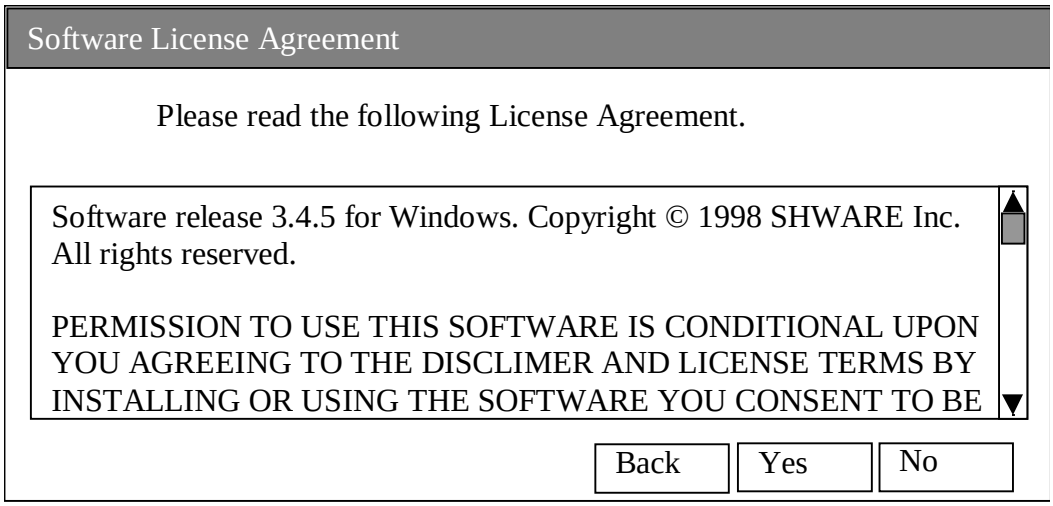

Figure 4: Displayed legal notice

2. An approach to a science of intent

The vision of intent information as metadata in a document raises the question of just what is this data, and how can it be represented. We could save the words or phases offered by the author as the reasons behind decisions, but this is qualitative, ad hoc and difficult to deal with. It is like saying an object should be red, but without the color-coordinates in a well-defined color space it is hard to know just what shade of red should be reproduced. For a science of intent we need things that can be measured. Measuring decisions directly does not seem best because a decision may be made for different reasons and every document has a different set of decisions. However, one can ask what is the reason behind each decision. The reasons given are typically to improve some property that is considered good. We call these value properties and they are things such as distinguishability, legibility, economy, balance, group identity, uniformity and how eye-catching it is. There are lots of possible value properties; every rule in a document design book is a potential property. Just as decisions occur at all stages of document processing, so do the value properties behind them. However, there is typically a hierarchy of properties as one moves from the specific to the general. For example, one may select a halftone in order to improve the smoothness of sharp edges, so that characters are easier to recognize, so that text is more legible, so that the document communicates more effectively. Value properties help to understand decisions, but do not capture intent directly. Intent is found in the relative importance of the value properties. We cannot maximize all of the value properties simultaneously. If documents could be designed to do this then we would just construct documents this way without having options or decisions. What is really behind our choices is an assignment of importance to certain value properties over others. In the phone book example, a small font is chosen because it reduces the amount of paper needed and improves the economy value property; but there is also the assignment of greater importance to economy than to other properties such as legibility. For the office phone list, legibility is more important than economy and so a decision that maximizes a different value property is made. The intent is then revealed in the relative weightings of the value properties. Let us assume that there are a handful of high-level intents that govern the decisions. The fact that there can be lots of value properties implies that the weightings of importance of value properties should be correlated. An approach to determining the space of intents is then to measure a large set of value properties for a diverse set of documents and perform a factor analysis to look for statistical correlations between the properties. The number of significant factors should provide the dimensionality of intent space. Factor scores could provide an estimate of a document's intent as a means of generating the intent metadata. The factor loadings matrix would tell how strongly each intent should weight each value property.

3. Issues

Work on just such an analysis is in progress, however, there are some issues that must be addressed for the approach outlined above. One is the completeness of the set of value properties to be examined. An intent factor will only be revealed by the correlation between value properties. If an intent does not have several measured value properties that are within its scope, then it will not be revealed. A possible way to check completeness might be to see whether the value properties, when weighted by

the intent factors, are sufficient to generate the desired decisions, but this in turn has issues such as the determinism of document decisions.

Another issue is how to define and implement any particular value property. Ideally, each value property would be based upon a set of experiment rating and ranking user responses with respect to the property over the various decisions. While some experimental data (such as letter size and legibility) exists, most value properties are found as qualitative design rules, relying on the judgement of the designer and lack quantitative expression. In the interest of time our solution is to use "best guess" approximations to the value properties and hope that the behavior is mimicked sufficiently to preserve the correlations.

A third issue is dealing with built-in correlations. Value properties can have correlations due to their definitions as well as the desires of the author; for example, cost, transmission time and information content may all be defined using the number of characters in the document and would therefore show a correlation. Such correlations can be very strong and in the factor analysis they can overwhelm weaker intent correlations.

4. Summary

Document intents provide useful information for use in document processing operations that are outside the immediate control of the author. This is particularly important for web documents that can be displayed on a variety of devices. An approach to quantifying intents is through the statistical analysis of document value properties. This approach, however, still has several issues that must be addressed.