A NEW APPROACH TO THE AUTOMATIC WEB ACCESSIBILITY

Juan Manuel Fernández[1], Vicenç Soler[1,2]

[1]Dept. Microelectrònica i Sistemes Electrònics, Universitat Autònoma de Barcelona, Cerdanyola del Vallés, Spain
[2]Ciber-BBN: Bioengineering, Biomaterials and Nanomedicine. Campus UAB, Bellaterra, Spain

Jordi Roig

Dept. Microelectrònica i Sistemes Electrònics, Universitat Autònoma de Barccelona, Cerdanyola del Vallés, Spain

Keywords: Web Accessibility, World Wide Web Consortium, Web Standardization, HTML, CSS.

Abstract: The number of Web pages created in an inaccessible manner is very high, and to convert them to be accessible can be impossible. We present a study of different tools that can help to obtain a total Web Accessibility. With this study, we can see the advantages of the automatic correction and the possibilities that this kind of tools offer. The result of the study shows how the applications adapt the Web pages to both the World Wide Web Consortium HTML grammar and to the Web Content Accessibility Guidelines. With this study we find an automatic way to improve Web Accessibility and to solve the problems that it can offer to the Web pages authors. This way consists of combining two tools. This combination gives us results that allow disabled people to use a system that can convert the entire Web in an accessible way.

1 INTRODUCTION

In this paper we present a study of different tools that can improve the Web Accessibility. With this analysis we try to find a tool, or a combination of tools, that can help the e-integration. The potential of the Web in the integration of disabled people is very high, but in order to exploit it, it is necessary that Web page developers use correct the grammar of the HTML (W3C, 1999b) and the accessibility guidelines (W3C, 1999b). Both of them are dictated by the World Wide Web Consortium (W3C). In the case of the accessibility, it exists a group of especial interesting, called the World Accessibility Initiative (WAI), inside of the W3C.

The use of the HTML grammar in the Web is very low. The lack of knowledge of this grammar, on the part of the Web page designers, has the origin in the fact that they design to a specific Web browser but these Web browsers are not standard (The Web Standards Project, 2007). If the use of the correct HTML grammar is very low, the use of the norms developed for disabled people is very poor. A great number of Web designers do not know the special needs of the disabled people.

The purpose of this study is to arrive to a strategy that can allow us to transform Web pages to the W3C and WAI normative. More specifically, the norms we are interesting in are the Web Content Accessibility Guidelines (W3C, 1999b) (WCAG). These guidelines mark what and how it is correct to present information by the Web. In this paper, we compare two tools that work in this manner.

Accessible Web is a tool that works with WCAG (ACTAW) (Fernández et al., 2007). This tool works with HTML, XHTML and CSS, and it applies the WCAG guidelines to a Web document.

The other tool that we use in this study is the HTML TIDY project (Dave Raggett, 2003). This application tries to adapt HTML and XHTML code of the Web page to the correct grammar.

The study compares the tools in two manners. First of all, we concentrate the study on the HTML grammar. Second, we analyze the results under the perspective of the Web Accessibility. With these studies we can try to find a possible way to improve Web Accessibility.

2 TOOLS USED IN THE TEST

2.1 ACTAW

The aim of the ACTAW is to improve Web Accessibility by applying the WACG to a Web page. The result of applying ACTAW to a Web page is a document with a better HTML structure and better accessibility. Moreover, ACTAW can modify the styles implemented with CSS and offer styles with a high colour contrast.

This software implements an HTML analyzer and a standardization module. The parser analyzes the HTML tags and attributes of a Web page, but it is not based on its structure or the correct W3C grammar. This fact helps to understand all kind of Web pages.

The standardization module has the objective of improving Web Accessibility. This module applies the WACG to the documents as well as it is possible. It is impossible to apply some guidelines, like the one to use an easy language.

2.2 HTML TIDY

HTML TIDY is an Open Source solution to obtain a correct HTML document. This tool transforms the HTML original source from a Web page to a correct HTML code that follows the W3C grammar, but its effects reach the Web Accessibility as well.

The document, after the execution of HTML TIDY, contains valid HTML tags and it corrects the structure of the document if it is not valid. It means that HTML TIDY can determine if the use of the elements of that markup are well positioned and replaces them by the correct solution.

2.3 W3C Validation Service

The first point of our interest is to measure the correctness of the Web pages results from ACTAW and HTML TIDY. To know if they are correct, it is necessary the W3C Validation Service (W3C, 2007a). This online tool analyzes the document and creates a report with the points where the HTML, or XHTML, grammar is not accomplished.

2.4 TAW Validation System

Test de Accesibilidad Web (TAW) (Fundación CTIC, 2007) is a tool developed by CTIC foundation, with the support of the Spanish Ministry of Education and Science. The Spanish W3C office forms part of this foundation.

The analyzer validates the different guidelines of WCAG 1.0. The result offers the number of errors organized in the three priorities, and the way the error can be validated. The report is very exhaustive and the summary offers a quick view of the problems found and if it needs the interaction of the designer.

2.5 Test Bench

The elements of study are the real Web pages. We use a group of sixty real Web pages selected randomly in an automatic way. This test bench is very heterogeneous and contains a different kind of contents. An interesting point is the fact that it contains Web pages designed in XHTML or HTML and that they can be correct or not.

3 PROCESS OF THE ANALYSIS

3.1 Study of the Test Bench

First of all, we need to know what kind of errors has every Web page to compare with the results of the process. To solve this point it is necessary an exhaustive study of the test bench with the focus on HTML grammar and the WACG.
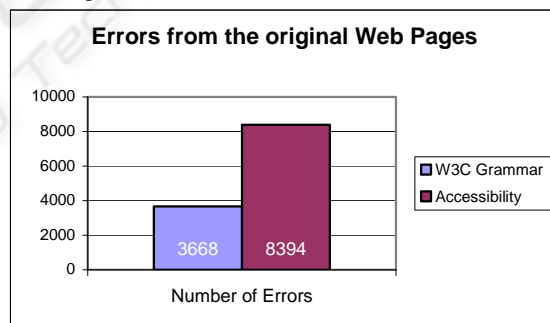

Figure 1: Errors from the Original Web Pages.

In Figure 1, we can see the totality of the grammatical and the accessibility errors. On average, it is the same those 23 grammatical errors by Web page, and 139.9 accessibility errors. This difference allows us to see the low knowledge of the Web Accessibility development.

The priority 2, which has guidelines affect the styles used, the creation of tables and scripts, is the biggest with difference (See Figure 2). These points are very difficult to accomplish if you do not have the knowledge of the correct manner of doing it. And also, in this priority we can find the correct use

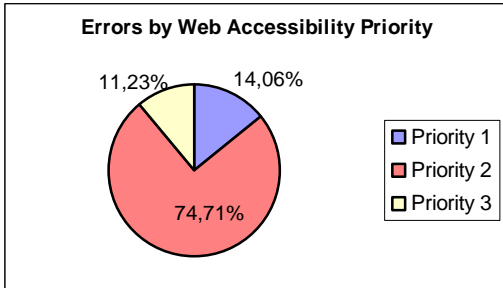of the HTML tags, in other words, the use of the correct grammar.

Figure 2: Errors by Accessibility Priority.

3.2 Applying HTML TIDY

The second step in the process is to apply HTML TIDY to the test bench. The improvement of the accuracy to the W3C grammar is very high (see Figure 3). With a 58.64% of improvement, HTML TIDY has very good results. This is the same that to have 11 errors by Web page on average. HTML TIDY deletes non-standards tags and attributes, one of its main problems, and it repairs structures which do not accomplish the standard.
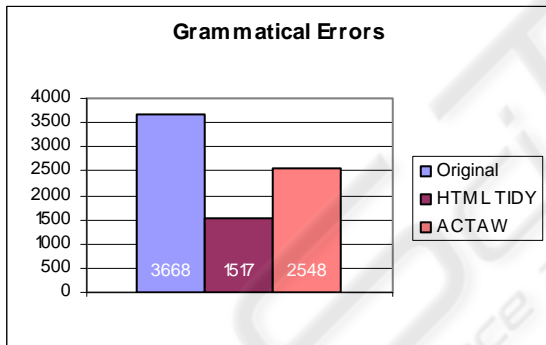

Figure 3: Grammatical Errors.

HTML TIDY does not try to solve accessibility errors, but it solves some inaccessible situations due to the final correct structure. It is surprising the fact that the best improvement is in the priority 1, which is 16.44%, and only 5.39% at the priority 2. The Figure 4 shows the results of applying HTML TIDY.
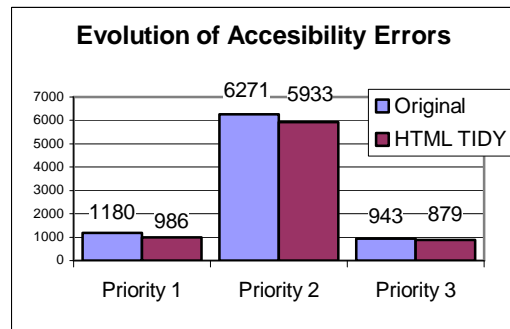

Figure 4: Evolution of Accessibility Errors.

3.3 ACTAW Test

ACTAW offers an improvement of the 30.53% of the correctness of the document in the test of the W3C without delete any tag (see Figure 3). We will discus about it in section 3.4.1.

Where ACTAW has excellent results is in the field of Web Accessibility. The tool obtains a 96.19% of improvement in the errors from priority 1. This is the same than having 0.75 errors by Web page. If we observe the results in depth we can see that the 71.67% of the Web pages do not have any priority 1 error and can obtain the level A in conforming with the WCAG. If we look the rest of the priority, we can see a reduction of the errors. An 86.85% of improvement is obtained for the priority 3 which is very impressive.
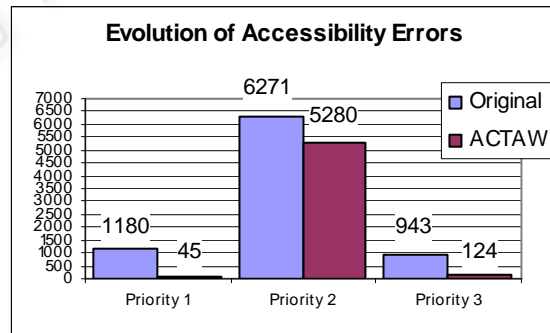

Figure 5: Evolution of Accessibility Errors.

Figure 5 shows the evolution of accessibility errors between the original Web pages and the results of ACTAW, and we can see that priority 2 has a low improvement comparing it with the rest of priorities. The reason of it is the complexity of the problems in this priority. We can emphasize two points: the design of the scripts and the use of the tables.

395

3.4 Comparative of the Results

3.4.1 HTML Grammar

The differences between the two tools are considerable, but it is necessary to comment one of them. ACTAW does not delete the tags which do not form part of the W3C grammar. In opposition to this, HTML TIDY deletes them. Both actions are correct but it is difficult to know which one is better.

The best action is to replace the incorrect tags by the equivalent ones. The existence of equivalent tags is not possible in all the cases and the number of non-standard tags is very high. The option to delete the tags and attributes can make the Web page not working properly. It is important if the Web page is a graphical interface of an online application.

The best solution for problems with incorrect tags and attributes is HTML TIDY. The process of validation shows a reduction of 2151 errors. In this field, ACTAW offers an improvement of the 30.53% (see Figure 2).

3.4.2 Web Accessibility

The results offered by the tools in this field are very good; in all the priorities they reduce the amount of errors. We have commented the good results of the tools, but now we are interested in priority 2. In this priority both tools cannot offer excellent results (see Figures 3 and 4), but ACTAW improve in an 11.01% the results of HTML TIDY. This improvement can be low but it is compensated by the results on the other priorities.

4 ACTAW & HTML TIDY

After comparing the different tools we are going to study the combination of both. The amount of grammatical errors shows us that the best combination is to correct grammar errors and after the accessibility errors. We reduce 2512 errors if we apply ACTAW and after it HTML TIDY. This is an improvement of the 73.34%. In comparison with the application of HTML TIDY alone, we have improved a 14.70%.

If we take the Web Accessibility errors we obtain the same results. It is not important the order of the application of the tools. The improvement of the errors is 32.95%, which is a very good result, but lower than the ACTAW results.

5 CONCLUSIONS

We have seen a comparative between two tools that work transforming automatically Web pages. Both of them try to apply the standards of the W3C. HTML TIDY works on the correction of the HTML grammar and ACTAW transforms Web pages to be accessible to disabled people.

The comparative shows us that the tools work correctly outside of its field and have great results on its field. After studying the tools, we have studied the result of their combinations. The results show that the combination of ACTAW and HTML TIDY can offer an improvement of the 73.34% in the correction of the HTML source. If we think in Web Accessibility, the combination can improve a 32.95%. This combination offers a very good improvement of the Web.

In conclusion, we can say that we have found a possible automatic solution to improve Web Accessibility. The combination of different tools can allow us to obtain a way to correct the low index of accessible Web pages.

REFERENCES

Fundación CTIC. "Test de Accessibilidad Web 3.0". http://www.tawdis.net/ (Retrieved on September 2007)

Dave Raggett, 2003. "TIDY Project". http://www.w3.org/People/Raggett/tidy/ (Retrieved on March 2007)

Juan Manuel Fernández, Vicenç Soler and Jordi Roig, 2007. Automatic Conversion Tool for Accessible Web. In Proceedings of the 3rd International *Conference on Web Information Systems and Technologies. (Webist'07)*

The Web Standards Project. http://www.webstandards.org/ (Retrieved on May 2007)

World Wide Web Consortium (W3C), 1999a. HTML 4.01 Specification http://www.w3.org/TR/html401/ (Retrieved on June 2007)

World Wide Web Consortium, The W3C Validator Team, 2007a. "The W3C Markup Validation Service". http://validator.w3.org/ (Retrieved on June 2007)

World Wide Web Consortium (W3C), 1999b. Web Content Accessibility Guidelines 1.0, http://www.w3.org/TR/WCAG10/ (Retrieved on June 2007)