

Web サイトのメニューの読み上げ順序の決定手法

津村 光輝† 岩田一†† 白銀 純子††† 深澤 良彰†
早稲田大学† 神奈川工科大学†† 東京女子大学†††

1 はじめに

インターネットの普及に伴い、Web ページの構造の複雑化やコンテンツの多様化が進んでいる。例えば、jQuery などの技術を用いて動的な変化をする Web ページや、Twitter のような SNS で一つの画面に膨大な情報が含まれているものがある。また、インターネットの普及に伴い、インターネットを利用する視覚障害者も増加している。電気通信普及財団による調査によると、全盲者の 96.3% はコンピューターを利用して、その主な利用目的は、Web 閲覧である [1]。Web 閲覧をする上で、視覚障害者の多くが利用するソフトウェアがスクリーンリーダーである。

現状のスクリーンリーダーは、訪問している Web ページの HTML ファイル内のテキストを基本的に上から下に読み上げるため、ユーザーが望まない箇所が読み上げられてしまうことがある。他にもメニュー読み上げが後に来ると、メニューから目的の Web ページを訪問するまでに時間がかかる。Lazar らは、スクリーンリーダーを使用しているユーザーは平均で 30.4% もの時間を無駄にしている [2] という。これは読み上げの順序や alt のない画像によるものであるという。

これらの問題に対し本研究では、Web サイトのメニューに着目し、メニューの読み上げ順序の決定手法を提案する。これにより効率的な Web ブラウジングが可能になると考える。

2 Web ページ構造

Web ページは一般的にグローバルナビゲーション、サブメニュー、コンテンツから成る。グローバルナビゲーションは Web サイトの全ての Web ページに共通した案内リンクのことを指し、サブメニューとは階層的に表示されるメニューのうちの下位のメニューのことを指す。コンテンツは Web ページで提供されている主な情報を指す。以下の図 1 は Web ページ構造の例である。黄色の部分グローバルナビゲーション、オレンジ色の部分がサブメニュー、青色の部分がコンテンツである。



図 1 Web ページ構造

3 提案手法

まず最初に、アクセスした Web ページからナビゲーション部分を抜き出し、次に、Web ページの各部分の読み上げ順序を決定する。まずはじめに、Web ページの HTML ファイルからそれぞれグローバルナビゲーション、サブメニュー、コンテンツを探し出し、その後、それぞれの部分の読み上げの順序を決定する。

3.1 ナビゲーション抜き出し

グローバルナビゲーションおよびサブメニューの抜き出しは次に示すアルゴリズムで行う。大まかなアルゴリズムとしては、まず始めに、与えられた URL から HTML を取得して、グローバルナビゲーション

A Method to Determine Reading Order of Website Menu
Koki Tsumura, Yoshiaki Fukazawa Waseda University †
Hajime Iwata Kanagawa Institute of Technology ††
Junko Shirogane Tokyo Woman's Christian University †††

およびサブメニューの候補として探索するタグ (nav タグか ul タグ) の内側にある HTML を以下の図 2 のように抜き出す。その後、それらの中から、footer タグで囲まれているものを削除し、探索対象のタグの内側にある HTML の内部の a タグが示す URL をそれぞれ訪問する。訪問先の URL にて一番最初に与えられた URL の内部にあったグローバルナビゲーションおよびサブメニューの候補があるかどうかを確認する。その候補があった場合には、その候補のカウントをインクリメントする。すべての a タグが示す URL を訪問後、候補のタグのカウントの中で、最大のものをグローバルナビゲーション、二番目に数の大きいものをサブメニューとする。



図 2 各部分の抜き出し

3.2 Web ページの読み上げ順序の決定

読み上げ順序は、グローバルナビゲーション、サブメニュー、コンテンツのすべてを読み上げるパターン 1 と、サブメニューとコンテンツを読み上げるパターン 2 と、コンテンツのみ読み上げるパターン 3 の 3 パターンである。パターン 1 は、最初に Web ページへ訪問したときに読み上げられるパターンであり、パターン 2 は前に訪問していたページから、Web サイト内の違うページに移動した際、サブメニューが前の Web ページと異なるときの読み上げパターンである。パターン 3 は現在のページとその前に訪問していた Web ページのサブメニューが同じ時の読み上げ方である。例えば、図 1 の Web ページに最初に訪問した際には、パターン 1 で読み上げられ、その Web ページから授業リンクをクリックして、図 3 の Web ページを訪問した際には、サブメニューが異なるのでパターン 2 で読み上げられる。また、図 1 の Web ページを訪問した後に、サブメニューのお知らせリンクから、図 4 の Web ページを訪問した際には、パターン 3 で読み上げられる。



図 3 パターン 2



図 4 パターン 3

4 おわりに

本研究では、Web ページのグローバルナビゲーション、サブメニュー、コンテンツに着目して、新たな Web ページの読み上げ順序の決定手法を提案した。今後の課題としては、以下のものがある。

- より良いナビゲーション抜き出しの検討 (Zhao らの研究 [3] などを検討)
- コンテンツ部分の読み上げ順序の決定

参考文献

- [1] 財団法人 電気通信普及財団. 2013. 「視覚障害者の携帯電話・スマートフォン・タブレット・パソコン利用状況調査」
- [2] Jonathan Lazar, Aaron Allen, Jason Kleinman and Chris Malarkey. 2007. "What Frustrates Screen Reader Users on the Web: A Study of 100 Blind Users", In proceedings of INTERNATIONAL JOURNAL OF HUMAN-COMPUTER INTERACTION, 22(3), pp247-269
- [3] Kui Zhao, Bangpeng Li, Zilun Peng, Jiajun Bu and Can Wang. 2017. "Navigation Objects Extraction for Better Content Structure Understanding". In Proceedings of WI ' 17, Leipzig, Germany, August 23-26, 2017