Measuring Disorientation Based on the Needleman-Wunsch Algorithm

Tolga Güyer, Bilal Atasoy, and Sibel Somyürek
Gazi University, Turkey

Abstract

This study offers a new method to measure navigation disorientation in web based systems which is powerful learning medium for distance and open education. The Needleman-Wunsch algorithm is used to measure disorientation in a more precise manner. The process combines theoretical and applied knowledge from two previously distinct research areas, disorientation and string-matching. String-matching algorithms provide a more convenient disorientation measurement than other techniques, in that they examine the similarity between an optimal path and learners' navigation paths. The algorithm particularly takes into account the contextual similarity between partly relevant web-pages in a user's navigation path and pages in an optimal path. This study focuses on the reasons and the required steps to use this algorithm for disorientation measurement. Examples of actual student activities and learning environment data are provided to illustrate the process.

Keywords: Student evaluation; evaluation methods; disorientation measurement; web navigation

Introduction

The Web is a powerful learning medium for distance and open education (Khan, 1997 p.5). On the one hand web based systems expand educational opportunities; on the other hand these systems cause specific problems (Hara, 2000). Disorientation is cited as one of two fundamental problems in the navigation of web-based environments (Conklin 1987; Hammond 1989; Beasley & Waugh 1995; Nielsen 1990; Shin et al. 1994; Mcdonald & Stevenson 1998; Chen 2002; Rezende & Barros 2008; Oostendorp & Juvina 2007; Hill & Hannafin 1997). In web systems, users must make decisions such as "where to go next, how to get there, and where they are in the overall structure of document" (Elm & Woods 1985; Zhu 1996, 26). Due to these responsibilities, most web users experience occasional difficulties when browsing and lose their ability to find their way within large-scale environments (Kim & Hirtle 1995; Dillon et al. 1990; Nielsen 1990).

Due to the users' ability to navigate freely, instead of following a previously created path, hypermedia is seen as an effective tool to deliver information to students (Alomyan, 2004). However, this freedom of navigation also causes several usability problems, as students need to make extra efforts to understand which content is relevant to their instructional objectives, and to develop a clear idea of the relations among the pages/links and their position in the environment (Zhu 1999, 1998; Brusilovsky & Henze 2007). While students attempt to conduct a task, at the same time they need to orient themselves in the environment. Managing these multiple tasks creates divisions in their mental processes, with the result that they pay less attention to their learning task (Tripp & Roby 1990; Kim & Hirtle 1995). In addition, missing some of the relevant content may decrease the students' learning performance (Mitchell, Chen, & Macredie, 2005). A large amount of research has been conducted to overcome or reduce these problems. Some studies have focused on providing navigation aids, such as a map, an index, guided tours, bookmarks, fish-eye views, history tools, page labels, or summary boxes (Parunak, 1989; Nielsen, 1990; Jonassen, 1992; Gupta & Gramopadhye, 1995; Kim & Hirtle, 1995; Ziefle & Bay, 2006; Pilgrim, 2012; Ruttun & Macredie, 2012). In a few of these studies, an attempt was made to personalize the system for each individual, and to adapt provided links according to the profile or needs of each user (Brusilovsky, 1998; Chen, 2002; Juvina & Herder, 2005; Hwang et al., 2010).

Other studies have been conducted to understand disorientation and several measurement techniques have been offered to define the degree of user disorientation more conveniently. In the long history of research on disorientation, the developed techniques can be classified under four main topics: performance, subjective opinions, metrics, and optimal path. These techniques are briefly explained below.

Performance: In early studies, it was thought that disorientation could be measured according to the performance of the users. The time spent to accomplish a specific task or the accuracy of answers was seen as a good predictor of performance. So, if the time spent is prolonged or if mistakes increase during the performance of tasks, the users are thought to be more disoriented (Edwards & Hardman, 1989; Yatim, 2002).

Subjective opinions: Some researchers have suggested new methods that are subjective in nature. These are used to gather information on the users' perceptions of disorientation (Ahuja & Webster, 2001; Beasley & Waugh, 1995). These measurement tools consist of several Likert-type scale questions. Although they are simple and quick to administer, they are limited in that the researcher may never know if the user is really aware of her/his disorientation.

Metrics: When users interact with a web-based environment, their navigation logs are stored in databases. Several metrics were generated from analyses of these navigation logs to measure disorientation in various studies (Herder & Juvina, 2004; Juvina & Oostendorp, 2006; Gwizdka & Spense, 2007). While some metrics were derived directly from the raw navigation data, others were calculated from weighted graph-based models of the hypermedia (Gwidzska & Spence, 2007; Juvina & Oostendorp, 2006). For example, Herder (2003) indicated that median view time and return rate are strongly associated with disorientation. Number of cycles is also seen as a good predictor of disorientation, according to a study conducted by Juvina and Oostendorp (2006).

Optimal path: Optimal path is the best way for users to navigate learning resources in order to achieve a specific learning goal. If the user utilizes the optimal path, s/he probably will not stray from the learning goal, and thus will not experience disorientation. The formulae developed by Smith (1996) and Dias and Sousa (1997) calculate disorientation based on the ratio of visited and optimal web pages.

Although there are some formulae that examine variances between users' navigation paths and optimal paths, similarity matching is considered a more useful and successful method to measure disorientation. Similarity matching is used to measure the similarities between two different strings (Chang & Lampe 1992). If a navigation path is conceived as a string consisting of page sequences, the operation actions needed to transform the user's navigation path to the optimal path can be determined by similarity matching algorithms.

The Needleman-Wunsch (N-W) algorithm was proposed in 1970 by Saul Needleman and Christian Wunsch. It is commonly used for general sequence alignment and scoring. Although the original purpose of this algorithm was to search for similarities between protein or nucleotide sequences in bioinformatics (Needleman & Wunsch, 1970), it is used far more widely today (Kozik 2011; Skopal & Bustos, 2011; Liner & Clapp, 2004; Gwizdka & Spense, 2007; Stefik, Hundhausen & Patterson, 2011). This string-matching algorithm, which includes four parameters, finds the optimal alignment of two sequences and computes similarities between them. The first two parameters in the algorithm are two strings, which ideally would match. The third parameter is a similarity matrix, showing relations between each character of the two strings. The fourth is a gap penalty which is a value designed to reduce the score when the characters do not match.

Disorientation is considered a serious and common navigation problem faced by students studying in web-based education environments. Since most of the distance education programs

are delivered with web systems, finding students' disorientation levels and preventing it is important to improve learning. For that reason, correct disorientation measurement is obvious and inevitable. To measure disorientation, a useful approach is to search for the similarities between the pages which students visit and the pages which are optimally related to her/his task. In several fields, such as bioinformatics, linguistics, and artificial intelligence, a variety of string-matching algorithms are used to find similarities. This study focuses on joining two previously distinct research areas, disorientation and string-matching algorithms, to increase the quality of distance education by using theoretical and applied knowledge from two disciplines.

The N-W algorithm provides a more precise disorientation measurement than other disorientation measurements, because it takes into account the contextual similarities between partly relevant pages in a user's navigation path and those in an optimal path. The algorithm provides the opportunity to find the common patterns present in learners' navigation logs and optimal path for the task. The traditional similarity view on disorientation measurements focused on exactly overlapping nodes in a user's navigation path and in an optimal path. However users' navigation trails can include partly relevant pages. With the matrix parameter in the algorithm, these partly relevant pages can be a part of measurement and different scoring schemes for the page assignments can be determined contextually. This study discusses reasons to use the N-W algorithm to measure disorientation. In addition, all of the steps for using the N-W algorithm are clearly explained, and examples of students' activities and learning environment data are provided for illustration and to guide further studies.

Method

This study is focused on development and use of a new method to measure disorientation. In detail, firstly a web-based learning environment was developed, and students' navigation paths were observed within this environment in an ICT course. Disorientation scores were initially computed using one of the most cited measurement methods developed by Dias and Sousa (1997) to examine users' paths compared to an optimal path. After calculating disorientation with this method, missed points were examined. Taking these shortcomings into account, the desirability of an improved method became apparent. Then, as an improved disorientation measurement method, implementation of the N-W algorithm was represented, the steps in the use of this algorithm are explained, and the method's results are illustrated.

The Web Based Learning Environment

The web-based learning environment developed for this study is focused on teaching HTML in an undergraduate information and communications technology (ICT) course. To provide a student's navigation behavior goal-directed, the learning system designed as task-based and five different tasks were constructed for five weeks experimentation process. The content of the learning

environment included seven main sections that covered core aspects of HTML: introducing HTML, basics in HTML, HTML texts, HTML images, HTML links, HTML tables and HTML frames, and constructed by fifty-three nodes. Instructional materials for each topic were created by videos in which the content was captured audio-visually. The videos firstly explain the subject theoretically (for example "What is frame?" or "Why we use it?") and then they show practically how they are used in real life situations by code examples. PHP, a server-side scripting language, was used to develop a dynamic web-based learning environment and MySQL database was used to store data about users, navigation logs, content, etc.

Participants

To collect navigation data of students and to perform validity of the new disorientation measurement method, 80 undergraduate students in a public university participated in the study. The participants ranged in age from 18 to 24, with a mode age of 19.6. Forty-two (52.50%) of the participants were male, and thirty-eight (47.50%) were female.

Needleman-Wunsch Algorithm

The Needleman-Wunsch (N-W) algorithm is a dynamic programming algorithm for optimal sequence alignment (Needleman and Wunsch, 1970). The N-W algorithm uses four parameters for computing the similarity score of two strings. These parameters are:

- The first string to match

- The second string to match

- Similarity matrix (which has "number of nodes" rows and "number of nodes" columns)

- A penalty gap value for the unmatched characters

The higher score obtained from the N-W algorithm shows the better similarity. In order to make scores comparable with each other, the scores should be normalized.

The Disorientation Formula of Dias and Sousa

One of the most cited formulae (Dias & Sousa, 1996) to measure disorientation based on students' navigation logs and optimal paths is the following:

Disorientation=the number of relevant nodes for the task in the student's navigation / the total number of visited nodes in the student's navigation

This ratio measures the degree of disorientation because it determines the precision of information retrieval (Dias & Sousa, 1997). The higher score obtained from this formula shows lesser disorientation. If the score equals to 0, it means that the user was completely lost in the hypermedia system.

The Non-Linear Media Disorientation Assessment Tool

To assess the students' perceived disorientation, the Non-Linear Media Disorientation Assessment Tool, which was developed by Beasley and Waugh (1995), was used. This tool is a 5-point Likert-type self-report scale with seven questions. The lowest score that can be obtained from the scale is 7, whereas the maximum score is 35. Lower scores on this scale indicate lower perceived levels of disorientation and higher scores indicate higher disorientation levels. Internal-consistency reliability (Cronbach's alpha) for ease of use was 0.87.

Procedure

This study was carried out in the undergraduate course of "Information and Communications Technology" with the participation of 80 students in a large public university during five weeks. The experiment was conducted in sessions at the computer lab of the university. The course is presented to the students via a standard web-browser. The web based learning material was introduced to the participants in the first week and was available for the students from then on, for an overall five week time period. When students logged into the system, they were provided the tasks of the week and were told to complete them within one course session. Then, they met the task and worked on the contents and examples in any sequence in the web based learning material which contained 53 nodes. Students' logs of interactions when browsing in the learning material are stored in the database to compute disorientation. After studying the content, students conducted the task and upload it to the system. In order to ensure the validity of the new disorientation measurement method, The Non-Linear Media Disorientation Assessment Tool was administered to participants after each task.

Measuring Disorientation According to Formula Developed by Dias And Sousa

To use Dias and Sousa's (D&S) formula to measure disorientation, the first step is determining relevant nodes (pages) in the content for each task to be completed. In this study, the relevant nodes for each task were identified according to expert opinions. The invited 12 participants were experts in the field of educational technology and were selected for their high level of expertise in teaching HTML. For the extent of consensus and validation of answers by experts, the Delphi technique was utilized. The techniques applied with the repeated individual questioning of the experts. After four rounds, the optimal nodes for each task are determined.

For example, the relevant nodes for the third week's task are shown in Table 1.

Table 1

Optimal Path for the Third Task

2, 3, 5, 6, 17, 26, 27, 28, 29, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44

To use this formula, secondly we should handle each student's navigation paths that are stored in the database. As an example, the nodes visited by one student for completion of the third week's specific task are shown in Table 2.

Table 2

A Student's Navigation Path

1, 1, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49

According to this formula and the common nodes shown in Table 3, the disorientation of the student whose navigation path was shown in Table 1 would be computed as:

The number of relevant nodes for the task in the student's navigation path = 10

The total number of visited nodes in the student's navigation path = 20

DisorientationD&S = 10/20 = 0.50

Table 3

The Common Nodes Between the Student's Navigation Path and the Optimal Path of the Sample Case

Navigation path:	**1,1,** 32, 33, 34, <u>35</u>, <u>36</u>, <u>37</u>, <u>38</u>, <u>39</u>, <u>40</u>, <u>41</u>, <u>42</u>, <u>43</u>, <u>44</u>, 45, 46, 47, 48, 49
Optimal path:	2,3, 5, 6, 17, 26, 27, 28, 29, <u>35</u>, <u>36</u>, <u>37</u>, <u>38</u>, <u>39</u>, <u>40</u>, <u>41</u>, <u>42</u>, <u>43</u>, <u>44</u>

Shortcomings of the disorientation formula developed by Dias and Sousa.

The disorientation formula developed by Dias and Sousa is useful, because it determines the relevant node rate for the task among all the visited nodes. According to the formula, if a student's navigation path is exactly the same as the optimal path, his disorientation would be computed as 1. On the other hand, if the student's navigation path has no relevant pages that are related to the task, then his disorientation would be computed as 0. Thus, the disorientation score falls between 0 and 1; if it is close to 0, then the student is more disoriented, and if it is close to 1, then the student is less disoriented.

However, there is an important missing point in this calculation. Any two compared pages are considered either similar or not in this measurement. Yet, different degrees of relevance / similarity can exist between two different educational contents and even it is usually exist. For this reason, instead of handling a page relevant or not, a value should be determined for its relevancy. Thus, any page's relevancy may have a continuous value between 0 and 1. This value can be identified and taken into consideration in disorientation measurement with N-W algorithm. In this context, the significance of the N-W Algorithm becomes apparent.

Using N-W Algorithm to Measure Disorientation

To use the N-W algorithm for disorientation measurement, we need four inputs of the algorithm. Two inputs of the algorithm, the first and second strings, are sequences of characters, which ideally should match. In our sample case, these two strings are sequences of pages in the student's navigation path and those in the optimal path.

The first string to match="1, 1, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 35, 45, 46, 47, 48, 49" (the student's navigation path)

The second string to match ="2,3,5,6,17,26,27,28,29,35,36,37,38,39,40,41,42,43,44" (optimal path for the third task)

The third parameter, the similarity matrix, showing similarities among the pages in the educational environment, gives the similarity score between each character included in both of the compared strings. Thus, the labels for the rows and columns should be these characters. In our example, the labels on both sides of the similarity matrix include all of the pages in the educational environment, numbered 1 - 53. For each match at the intersection of every row and column, a similarity score typically between 0 and 1 is provided. These scores show the similarity of the pages to each other. This is of great importance, because the pages do not have to be completely similar or not. The scores are determined according to experts' opinions. An illustration of the similarity matrix is shown in Table 4.

Table 4

Similarity Matrix

	1	2	3	4	5	6	7	8	9	10	.	.	53
1	1	0.1	0.3	0.2	0	0.1	0	0	0	0	.	.	0
2	0.1	1	0.1	0.1	0	0.1	0.1	0.3	0	0	.	.	0
3	0.3	0.1	1	0.5	0	0.2	0	0	0.2	0.2	.	.	0
4	0.2	0.1	0.5	1	0	0.1	0	0	0.1	0.1	.	.	0
5	0	0	0	0	1	0.1	0	0	0	0	.	.	0
6	0.1	0.1	0.2	0.1	0.1	1	0.1	0	0	0	.	.	0
7	0	0.1	0	0	0	0.1	1	0.7	0	0	.	.	0
8	0	0.3	0	0	0	0	0.7	1	0	0	.	.	0
9	0	0	0.2	0.1	0	0	0	0	1	0	.	.	0
10	0	0	0.2	0.1	0	0	0	0	0	1	.	.	0
.	.	.	.	.	.	.	.	.	.	.	.	.	0
.	.	.	.	.	.	.	.	.	.	.	.	.	0
53	0	0	0	0	0	0	0	0	0	0	.	.	1

Several alignments between the student's navigation path and the optimal path can be seen. Two of the alignment examples for the sample case are shown in Tables 5 and 6. There, Op is the optimal path for the third task, and Np is the navigation path of the learner.

Table 5

An Alignment Example for the Sample Case

Op	2	3	5	6	17	26	27	28	29	35	36	37	38	39	40	41	42	43	44	_	_	_	_	_
Np	1	1	32	33	34	35	36	37	38	39	40	41	42	43	35	45	46	47	48	49	_	_	_	_

Table 6

Another Alignment Example for the Sample Case

Op	_	_	_	_	_	2	3	5	6	17	26	27	28	29	35	36	37	38	39	40	41	42	43	44
Np	_	_	_	_	1	1	32	33	34	35	36	37	38	39	40	41	42	43	35	45	46	47	48	49

The matching score between the student's navigation path and the optimal path must be computed consistently. While the N-W algorithm finds the highest score for each alignment, it will produce a standard score for each learner. For example, the alignment with the highest score should be as follows for our example.

Table 7

Alignment with the Highest Score between Navigation Path and Optimal Path in the Sample Case

Op	2	3	5	6	17	26	27	28	29	35	36	37	38	39	40	41	42	43	_	44	_	_	_	_
Np	_	1	_	1	_	_	32	33	34	35	36	37	38	39	40	41	42	43	35	45	46	47	48	49

The fourth parameter, gap penalty, can be scored as zero or a non-zero value depending on the situation. This score is used for each character of the sequence that is not placed alongside a corresponding character in the other sequence, and should be less than the minimum value obtained from the correct matches. This kind of match is illustrated using the "_" symbol in the matching tables.

The gap penalty is important in measuring disorientation, because it allows the researcher to distinguish a user performing a given task who navigated to the same relevant pages as another user in the same task who also navigated to irrelevant pages. So, the penalty gap must be determined as a non-zero score for disorientation measurement. The gap penalty value is determined as -0.1 in our study.

According to the best alignment shown in Table 7 and a gap penalty of -0.1, the N-W Algorithm would produce the following scores depending on the similarity matrix.

Table **8**

Scores for the Matching between Navigation Path and Optimal Path in the Sample Case

Op	2	3	5	6	17	26	27	28	29	35	36	37	38	39	40	41	42	43	_	44	_	_	_	_
Np	_	1	_	1	_	_	32	33	34	35	36	37	38	39	40	41	42	43	35	45	46	47	48	49
S	-0,1	0,3	-0,1	0,1	-0,1	-0,1	0	0	0	1	1	1	1	1	1	1	1	1	-0,1	0,1	-0,1	-0,1	-0,1	-0,1

Using the scores from the table, the disorientation score of our learner would be computed as follows:

Disorientation N-W = -0.1 + 0.3 - 0.1 + 0.1 -0.2 + 0 + 0 + 0 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 - 0.1 + 0.1 - 0.4 = 8.6

Standardized variables are needed to compare the student's disorientation. For that reason, the computed disorientation scores according to the N-W algorithm should be normalized by performing the following steps respectively:

1. The maximum possible score that can be produced by the algorithm for each task is computed. This score should be generated whether the learner's navigation path is exactly the same as the optimal path, and it should be equal to the number of nodes in the optimal path.

MaxScore = Number of nodes in the optimum path

In our sample case, this score would be 19.

2. The minimum possible score that can be produced by the algorithm for each task is computed. This score should be generated whether the learner's navigation path is exactly different from the

optimal path, and it should include all the irrelevant nodes. While computing this score, the number of irrelevant nodes should be multiplied by the gap penalty.

MinScore = (Total number of nodes in the educational environment - Number of nodes in the optimum path) x gap

For the sample case, the minimum score would be computed as follows:

MinScore = (53 - 19) x (-0.1) = -3.4

3. According to the maximum and minimum scores, each of the students' disorientation scores are normalized with the following formula. These scores fall between 0 and 1.

$$\text{Disorientation}_{N\&W}^{N} = 1 - \frac{\text{MaxScore} - \text{Disorientation}_{N\&W}}{\text{MaxScore} - \text{MinScore}}$$

Using the normalized disorientation formula, a higher disorientation score should be computed for a more disorientated student and a lower score for a less disorientated student. This is the reason for the subtraction from 1 in the formula.

If we return to our example, the normalized disorientation score of the learner is computed as follows:

$$\text{Disorientation}_{N\&W}^{N} = 1 - \frac{19 - 8.6}{19 - (-3.4)} = 0.53$$

As can be seen, the N-W algorithm disorientation score for the sample student's navigation path for the same task differs from the score generated by the method offered by Dias and Sousa. This difference derives from the method of determination of the similarities among the pages more precisely in the N-W algorithm than simply similar or not.

Validity and Reliability

The N-W algorithm is a commonly used algorithm to align two sequences. Its validity is examined and it is handled reliable as the optimal matching algorithm and applied for similarity analyses in several areas such as image processing, programming and seismic traces (Kozik, 2011; Skopal & Bustos, 2011; Liner & Clapp, 2004; Gwizdka & Spense, 2007; Stefik, Hundhausen & Patterson, 2011). For direct measurement of disorientation, an accepted objective measurement methodology is searching for similarities between students' navigation path and optimal path. In this context, to measure similarity between user's navigation path and pages in an optimal path, using N-W algorithm method is obviously valid. Also, the results of the Pearson Multiplication Moments Correlation Coefficient (r=72, p<.005) demonstrate that there is a positive relationship in the students' perceived disorientation scores according to the Non-Linear Media Disorientation Assessment Tool and students disorientation scores according to the new offered method based on the N–W algorithm.

The Main Steps of the Method

The seven basic steps and requirements needed to use the offered method to measure a learner's disorientation can be briefly summarized as follows.

Step-1: The educational environment should include an optimal path, and can be developed as a task-based and/or problem-based environment.

Step-2: The optimal path to achieve a given task should be decided. To determine relevant nodes, expert opinions can be used, as in this study; alternatively, task-relevance of pages may be automatically computed with semantic web technologies.

Step-3: The similarities among all nodes should be identified, and a similarity matrix should be constructed. For this stage, semantic web technologies also can be used.

Step-4: The learners' navigation data should be recorded.

Step-5: A penalty gap value should be generated for each irrelevant node (page) visit.

Step-6: While using the N-W algorithm, disorientation scores should be computed with the following inputs, students' navigation data, the optimal path, penalty gap value, and similarity matrix.

Step-7: The N-W algorithm disorientation scores should be normalized between 0 and 1.

Study Limitations

The offered method for measuring disorientation is developed for task based educational web sites. For each task, a similarity matrix is computed that shows the relevancy of any node (page) with tasks. However, a web page also can be designed as including different parts of information and every part can have different relevancy with task. This approach is limited for considering a single relevancy score for each page.

Discussion and Conclusion

Disorientation is considered a serious and common navigation problem faced by students studying in web-based education environments. Most learners experience difficulties concerning navigation that cause them to pay less attention to their learning tasks or to miss some relevant content. As a result, their learning performance is diminished. Several measurement techniques have been offered to define disorientation more conveniently. Among the variety of measures, analysing students' navigation paths is particularly useful, as this provides objective data, unlike Likert-type scales that are designed to gather subjective information from users about their perceptions of disorientation.

In this paper, we provided a new method to measure disorientation based on the N-W algorithm. With this method, researchers can search for similarities between students' navigation paths and an optimal path. This is a more accurate and precise measurement technique for two main reasons. First, it uses the actual navigation data as objective measurements, and it examines similarities between the pages that students visit and pages which are optimally related to the required task. In addition, this technique determines the best alignment between the user's and the optimal page sequences, and calculates the highest score for this alignment. Secondly, existing optimal path related formulae reduce the contextual similarity of pages in an educational environment to two distinct characterizations: similar or not. Yet, differing degrees of relevance/similarity between pages do exist. Degrees of similarity can be determined by expert opinions and can be scored using a continuous variable. The N-W algorithm utilizes various similarity scores between compared pages, instead of considering them either similar or not. Finally, determining an appropriate gap penalty score makes the N-W algorithm even more accurate, as the technique can distinguish between two users who navigate along the same relevant pages, when one also navigates to irrelevant pages (thus becoming more disoriented) but the other does not.

In the literature, only a few published studies have been conducted on the use of the N-W algorithm to measure disorientation (Gwidzska & Spence, 2005, 2007). These studies employed 10 different navigation factors to find predictors of disorientation and task success, including two factors named LCSMax and LCSLenMax. Although these two factors were derived from the N-W algorithm, their selected penalty gap score and similarity matrix construction did not facilitate a clear explanation of the process. To understand the N-W algorithm as a disorientation measurement method, the four inputs – user path, optimal path, similarity matrix and penalty gap scores – must be explained explicitly, with examples for illustration. This study was designed to provide such a guide for the use of the N-W algorithm.

In future studies, when using this method to measure learners' disorientation, the task relevancy of pages may be automatically determined by semantic web technologies instead of being decided by expert opinions. That should make the measurement process easier for researchers and increase its practical value, especially in large-scale environments. In this way, this technique can also be used in open corpus educational systems, which can constantly change and expand beyond the control of researcher. This study focused on disorientation measurements using the N-W algorithm. Lastly, this method can be used in adaptive hypermedia systems in order to identify students' disorientation in real-time, based upon the students' behaviors. To correct any discovered disorientation, navigation adaptations can be provided to guide users to the most appropriate pages for their tasks.

References

Ahuja, J. S., & Webster, J. (2001). Perceived disorientation: An examination of a new measure to assess web design effectiveness. *Interacting with Computers, 14*(1), 15-29.

Alomyan, H. (2004). Individual differences: Implications for web-based learning design. *International Education, 4*(4), 188-196.

Beasley, R. & Waugh, M. (1995). Cognitive mapping architectures and hypermedia disorientation: An empirical study. *Journal of Educational Multimedia and Hypermedia, 4*(2), 239-255.

Brusilovsky, P. (1998). Methods and techniques of adaptive hypermedia. In *Adaptive hypertext and hypermedia*, edited by P. Brusilovsky, A. Kobsa and J. Vassileva, 1-44. Boston: Kluwer Academic Publishers.

Brusilovsky, P. & Henze, N. (2007). Open corpus adaptive educational hypermedia. In *The adaptive web: Methods and strategies of web personalization*, edited by P. Brusilovsky, A. Kobsa and W. Neidl, 671-696. Lecture Notes in Computer Science, Vol. 4321, Berlin Heidelberg New York: Springer-Verlag.

Chang, W., & Lampe, J. (1992). Theoretical and empirical comparisons of approximate string matching algorithms. In *Combinatorial Pattern Matching*, 175-184. Springer Berlin:Heidelberg.

Chen, S. Y. (2002). A cognitive model for non-linear learning in hypermedia programs. *British Journal of Educational Technology, 33*(4), 449-460.

Conklin, J. (1987). Hypertext: An introduction and survey. *Computer, 20*(9), 17-41.

Edwards, D. M. & Hardman, L. (1989). Lost in hyperspace: Cognitive mapping and navigation in a hypertext environment. In *Hypertext: Theory into practice*, edited by R. McALEESE 105-125. Oxford: BSP Intellect books.

Dias, P., & Sousa, A. P. (1997). Understanding navigation and disorientation in hypermedia learning environments. *Journal of Educational Multimedia and Hypermedia.*

Dillon, A., McKnight, C., & Richardson, J. (1990). Navigation in hypertext: A critical review of the concept. Paper presented at the *Proceedings of INTERACT: The third IFIP Conference on Human Computer Interaction.* UK. London: Springer-Verlag.

Elm, W. C., & Woods, D. D. (1985). Getting lost: A case study in interface design. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Gupta, M. and Gramopadhye, K. (1995). An evaluation of different navigational tools in using hypertext. *Computers and Industrial Engineering*, *29*(1), 437-41.

Gwizdka, J., & Spence, I. (2005). What can searching behavior tell us about the difficulty of information tasks? Paper presented at the Special interest tracks and posters of the 14th international conference on World Wide Web (WWW '05).

Gwizdka, J. & Spense, I. (2007). Implicit measures of lostness and success in web navigation. *Interacting with Computers*, *19*(3), 357-369.

Hammond, N. (1989). Hypermedia and learning: Who guides whom? In *Computer Assisted Learning. Lecture Notes in Computer Science,* edited by H. Maurer, 167-181. Berlin: Springer-Verlag.

Hara, N. (2000). Student distress in a web-based distance education course. *Information, Communication & Society*, *3*(4), 557-579.

Herder, E. (2003). Revisitation patterns and disorientation. In German Workshop on Adaptivity and User Modeling in Interactive Systems ABIS03, 291-294. Karlsruhe. Germany: University of Karlsruhe.

Herder, E. and Juvina, I. (2004). Discovery of individual navigation styles. In *Proceedings of Workshop on Individual Differences in Adaptive Hypermedia at Adaptive Hypermedia 2004,* 40-49. Eindhoven, The Netherlands.

Hill, J. R. & Hannafin, M. J.(1997). Cognitive strategies and learning from the world wide web. *Educational Technology Research and Development*, *45*(4), 37-64.

Hwang, G.-J., Kuo, F.-Y., Yin, P.-Y. & Chuang, K.-H. (2010). A heuristic algorithm for planning personalized learning paths for context-aware ubiquitous learning. *Computers & Education*, *54*(2), 404–415.

Jonassen, D. (1992). Designing hypertext for learning. In *New directions in educational technology*, edited by E. Scanlon and T. O'Shea, 123-130. Berlin: Springer-Verlag.

Juvina, I. & Herder, E. (2005). The impact of link suggestions on user navigation and user perception In *Proceedings of the User Modeling 2005: Tenth International Conference (UM2005)*, Edinburgh, United Kingdom.

Juvina, I. & Oostendorp, H. V. (2006). Individual differences and behavioral metrics involved in modeling web navigation. *Universal Access in Information Society*, *4*(3), 258–269.

Khan, B. H. (Ed.). (1997). Web-based instruction. Educational Technology.

Kim, H. & Hirtle, S. C. (1995). Spatial metaphors and disorientation in hypertext browsing. *Behaviour & Information Technology*, *14*(4), 239-250.

Kozik, R. (2011). Improving depth map quality with Markov random fields. In *Image processing and communications challenges 3* 149-156. Advances in Intelligent and Soft Computing, 102, Springer: Berlin Heidelberg.

Liner, C. L. & Clapp, R. G. (2004). Nonlinear pairwise alignment of seismic traces. *Geophysics*, *69*(6), 1552-1559.

Mcdonald, S. & Stevenson, R. J. (1998). The effects of text structure and prior knowledge on navigation in hypertext. *Human Factors*, *40*(1), 18-27.

Mitchell, T.J.F., Chen, S.Y. & Macredie, R. D. (2005). Cognitive styles and adaptive web-based learning. *Psychology of Education Review*, *29*(1), 34-42.

Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology*, *48*(3), 443–453.

Nielsen, J. (1990). Navigation through hypertext. *Communications of the ACM, 33*(3), 297-310.

van Dyke Parunak, H. (1989). Hypermedia topologies and user navigation. In *Proceedings of the second annual ACM conference on Hypertext*, 43-50. ACM.

Oostendorp, H. v. & Juvina, I. (2007). Using a cognitive model to generate web navigation support. *International Journal of Human-Computer Studies*, *65*(10), 887–897.

Pilgrim, C. J. (2012). Website navigation tools – a decade of design trends 2002 to 2011. *Proceedings of the Thirteenth Australasian User Interface Conference (AUIC2012)*, Melbourne: Australia.

Rezende, F. & Barros, S. D. S. (2008). Students' navigation patterns in the interaction with a mechanics hypermedia program. *Computers & Education*, *50*(4), 1370–1382.

Ruttun, R. D. & Macredie, R. D. (2012). The effects of individual differences and visual instructional aids on disorientation, learning performance and attitudes in a hypermedia learning system. *Computers in Human Behavior*, *28*(6), 2182–2198.

Shin, E., Schallert, D. & Savenye, C. (1994). Effects of learner control, advisement, and prior knowledge on young students' learning in a hypermedia environment. *Educational Technology Research and Development*, *42*(1), 33-46.

Skopal, T. & Bustos, B. (2011). On nonmetric similarity search problems in complex domains. *ACM Computing Surveys, 43*(4), 1-50.

Smith, P. A. 1996. Towards a practical measure of hypertext usability. *Interacting with Computers, 8*(4), 365-381.

Stefik, A. Hundhausen, C. & Patterson, R. (2011). An empirical investigation into the design of auditory cues to enhance computer program comprehension. *International Journal of Human-Computer Studies, 69*(12), 820-838.

Tripp, S.D. & Robby, W. (1990). Orientation and disorientation in a hypertext lexicon. *Journal of Computer-based Instruction, 17*(4), 120-124.

Yatim, N. F. (2002). A combination measurement for studying disorientation. In System Sciences, 2002. HICSS. *Proceedings of the 35th Annual Hawaii International Conference 7.* IEEE.

Zhu, E. (1999). Hypermedia interface design: The effects of number of links and granularity of nodes. *Journal of Educational Multimedia and Hypermedia, 8*(3), 331-58.

Ziefle, M. & Bay, S. (2006). How to overcome disorientation in mobile phone menus: A comparison of two different types of navigation aids. *Human–Computer Interaction, 21*(4), 393–433.