Studying from Electronic Textbooks

Rakesh Agrawal Sreenivas Gollapudi Anitha Kannan Krishnaram Kenthapadi
Search Labs, Microsoft Research
Mountain View, CA, USA
{rakesha, sreenig, ankannan, krisken}@microsoft.com

ABSTRACT

We present *study navigator*, an algorithmically-generated aid for enhancing the experience of studying from electronic textbooks. The study navigator for a section of the book consists of helpful *concept references* for understanding this section. Each concept reference is a pair consisting of a concept phrase explained elsewhere and the link to the section in which it has been explained. We propose a novel reader model for textbooks and an algorithm for generating the study navigator based on this model. We also present the results of an extensive user study that demonstrates the efficacy of the proposed system across textbooks on different subjects from different grades.

Categories and Subject Descriptors

H.3 [**Information Storage and Retrieval**]: Information Search and Retrieval

General Terms

Algorithms, Experimentation, Human Factors

Keywords

Data Mining, Education, Electronic Textbooks, Study Navigator, Reader Model

1. INTRODUCTION

With the emergence of abundant online content, cloud computing, and electronic reading devices, the multi-billion dollar textbook industry is poised for transformative changes. Notwithstanding understandable misgivings (e.g. Gutenberg Elegies [7]), textbooks cannot escape what Walter Ong calls "the technologizing of the word" [18]. Already, there are initiatives such as "no child left offline" that are centered around the availability of electronic textbooks for achieving the goal of "any time, any place, any pace" learning [2]. These trends are not limited to USA or other developed nations. For example, Government of India is said to be developing a low cost tablet, Aakash, pre-loaded with educational content for distributing to millions of students [3].

CIKM'13, Oct. 27–Nov. 1, 2013, San Francisco, CA, USA.
Copyright 2013 ACM 978-1-4503-2263-8/13/10 ...$15.00.
http://dx.doi.org/10.1145/2505515.2505604.

We believe electronic textbooks provide huge opportunity to invent new tools and techniques to facilitate effective use of this medium. Some of the new functionalities that can be enabled in future textbooks include:

New navigations. The book can infer navigational aids beyond table of contents, back-of-the-book index, and simple hyperlinks.

Adaptability. The book can be personalized to suit the student's knowledge of the subject material as well as learning styles. The presentation can be dynamically modified to adapt to the requirements of the student based on prior interaction with the book.

Richer experiences. The book can be augmented with images, picture galleries, videos, and live simulations to provide a better learning experience.

Continuous self-assessment. The book can offer personalized assessments and recourses in a non-invasive way.

Collaborative learning. The book can propose interactions with other students, appropriate for the part of the book a student is studying. It can also suggest compatible study groups that can span different geographical regions.

This paper presents one particular study aid we have designed specifically for electronic books, called *study navigator*. The goal of the study navigator is to provide easy access to concepts explained elsewhere in the book that are most relevant for understanding the present section. Refer to the pair consisting of a concept useful for understanding a section and the link to the section where it has been explained as a *concept reference*. The study navigator consists of significant concept references for every section in the book. It can be activated by a student while reading a particular section and shows the corresponding concept references. Only a small number of significant concept references are shown to avoid undue cognition burden on the reader.

Our main technical contribution is the algorithmic mining of the concept references in the context of a reader model we propose for textbooks. Our reader model is inspired by the random web surfer model and personalized PageRank computation. However, the random walk used in our model has significant differences, such as (1) the preference vector gets updated during certain types of transitions while it is fixed in personalized PageRank computation, and (2) return transition occurs with a large probability in our model unlike in PageRank computations [5]. We also present the results of an extensive user study showing that the judges found the references provided by the study navigator to be quite helpful.

2. RELATED WORK

Authoring tools for adaptive navigation and presentation: A prominent system in this category is InterBook [8], a tool for creating an electronic book that can adapt to users with different backgrounds, prior subject knowledge, and learning goals. The data required to enable this adaptation must be provided as input by the author. We aim to infer the concept references needed for building the study navigator by algorithmically mining the text of the book.

Adaptive educational hypermedia systems: The goal of these systems is to combine hypermedia systems with Intelligent Tutoring Systems to adapt web-based educational material to the needs of particular users. They aim to help educators manually setup personalized courseware based on the cognitive style (*e.g.*, AES-CS [24], EDUCE [15]) or learning style (*e.g.*, KBS Hyperbook [13], INSPIRE [19]). They operate under the premise that the underlying information to enable this personalization is available to the person creating the courseware. We, on the other hand, aim to provide automated techniques.

Exploratory hypermedia systems: We put various systems categorized as ASK systems (*e.g.*, Trans-ASK [6] and ASKTool [11]) into this category. They aim to provide an interactive environment that mimicks conversing with an expert for its users to be able to find content of interest and/or ask follow-up questions to retrieve additional topics. In contrast, we look at the problem of identifying sections that are needed for understanding the current section.

Text browser and search tools: SuperBook [20], ScentIndex [10], ScentHighlights [9], and Smart [22] are examples of such systems. In contrast to our system, these systems do not provide references to concepts/sections that are useful for understanding a given section.

Back-of-the-book indices: While related, there are fundamental differences between a back-of-the-book index [17] and what we call concept references associated with each section of the book. In principle, one could do a sort on section numbers of a back-of-the-book index and thus find the important phrases present in each section. But it solves only half of the problem – if we know that a concept phrase φ is important for understanding a given section, we can use this approach to know all the sections where φ is possibly explained. But how do we know which φ is critical for understanding the present section? In fact, it is quite likely that φ might not even appear as a phrase in the present section. For the same reason, hyperlinking some phrases appearing in the current section is not sufficient. Another key difference is that back-of-the-book index generation algorithms compute global significance of concept phrases at the book level without taking into account where in the book the reader currently is.

3. STUDY NAVIGATOR

The study navigator system is designed to make it easy for a student to find concepts described elsewhere in the book that are most relevant to the material discussed in the present section. We refer to the pair consisting of a concept useful for understanding a section and the link to the section where it has been explained as a *concept reference*. For the purposes of this paper, we represent a concept as a phrase, and denote it as *cphr*. Our goal is to determine a few $\langle$*cphr*, section$\rangle$ pairs that are most relevant for understanding the current section.

3.1 Algorithmic Intuition

Suppose that the set of *cphrs* contained in a section as well as the relationship between *cphrs* is available[1]. We then need to determine the concept references that are most significant for understanding

[1]While multiple approaches are possible for computing *cphrs* and relationships between them [14], our implementation uses the approach provided in [4] for this purpose.

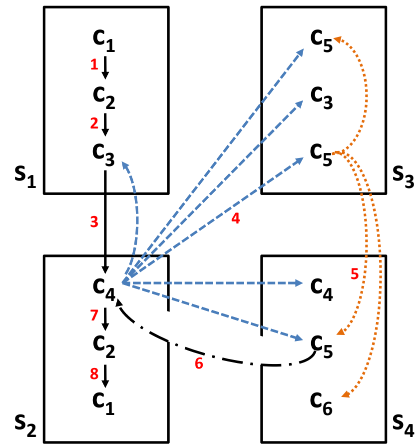


Figure 1: Illustration of Reader Model: Consider a hypothetical textbook consisting of four sections ($s_1, \ldots, s_4$) and six *cphrs* ($c_1, \ldots, c_6$). The reader reads the book starting from s_1. The path followed by the reader is indicated by numbers next to the arrows. Suppose the reader (after reading s_1) does not understand *cphr* c_4 in section s_2, and hence is forced to refer to another section containing c_4 or a *cphr* related to c_4. Let $\{c_3, c_4, c_5\}$ be the set of *cphrs* related to c_4, so that the available digression edges correspond to the edges consisting of dashes. The reader chooses a *cphr* from this set. Suppose she chose c_5. Out of the three occurrences of c_5 in the book, suppose she selected the second occurrence of c_5 in s_3. Thus she follows the digression edge marked 4, to read about c_5 in s_3. After reading about c_5 in s_3, the reader either returns to c_4 in s_2 with a large probability (the return edge not shown) or digresses further. Suppose she digresses further. Let $\{c_5, c_6\}$ be the set of *cphrs* related to c_5, so that the available digression edges correspond to the edges consisting of dots. She selects c_5 from this set and follows the digression edge marked 5 to read about c_5 in s_4. Afterwards, she returns to c_4 in s_2 along the edge marked 6, and persists to read further.

a given section s. For this purpose, we need a score denoting how significant is the description of a *cphr* c in a different section t for understanding section s. Given the significance scores of every *cphr* in every other section for understanding section s, we can order $\langle$*cphr*, section$\rangle$ pairs by their significance scores and include pointers to the top k $\langle$*cphr*, section$\rangle$ pairs in the study navigator for section s.

The significance score of a *cphr* in section t for understanding a different section s can be thought of in terms of how likely is the description of this *cphr* in section t to be referred when a reader is trying to understand section s. How do we formalize and quantify this likelihood? We surmise that while reading a book, a reader would refer to more significant *cphrs* more often.

Reader Model: Consider a student who is reading a textbook starting from the first section. When she is reading a section i, she comes across the *cphrs* in the order $c_{i1}, c_{i2}, c_{i3}, \ldots$. When the reader comes across a *cphr* c, with a large probability, the reader will be persistent in continuing to read the section. With a certain probability, she may not understand the *cphr* and hence may be forced to refer to another section to seek explanation.

Postulate that whenever the reader does not understand c, she refers to a section containing the same *cphr* c or a different *cphr* related to c. More precisely, the reader picks a *cphr* c' from the set of *cphrs* related to c with equal probability, chooses an occurrence of c' amongst all occurrences of c' in the book with equal probability, and refers to the corresponding section i' to learn more about c'. It is possible that i' is a section earlier than i in the book or it is a later section. After reading about c' in i', the reader has the following options: (a) return to the original section i with a large probability, and continue further reading, or (b) digress further to learn more

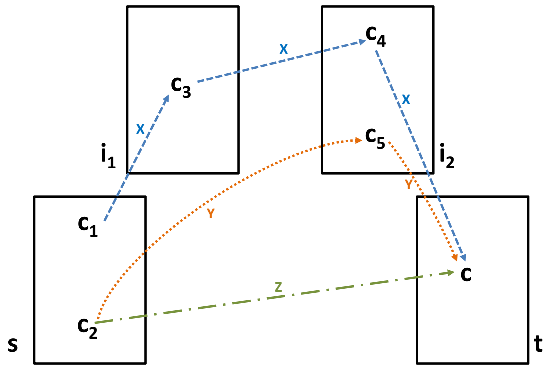


Figure 2: Illustration of how the significance score of *cphr* c in section t for understanding section s is computed: Consider three different readers trying to understand *cphrs* in section s. Reader X is unable to understand *cphr* c_1, and hence digresses to other sections (shown using dashed edges). She may first refer to c_3 in section i_1, followed by c_4 in section i_2, and finally c in section t. Readers Y and Z are unable to understand *cphr* c_2, but digress to different sections. Reader Y refers to c_5 in section i_2, followed by c in section t (shown using dotted edges) while reader Z directly digresses to c in section t (along the edge consisting of dashes and dots). The significance score is obtained by computing the likelihood of each such digression for different readers that reach c in section t starting from section s, and aggregating over many such digressions.

about c' by referring to a section containing c' or a different *cphr* related to c', that is, pick a *cphr* c'' from the set of *cphrs* related to c' with equal probability and refer to a section i'' that contains c'' amongst all occurrences of c'' with equal probability. In the latter case, the reader then returns to the original section i, or digresses further. Note that, while digressing, the reader can revisit a section i' (*e.g.*, for reading about c''' which is also explained in section i' and which is related to c''). But the return from a digression is always to the starting section i (irrespective of the number of hops digressed) as the reader is trying to understand section i and the purpose of the digression is to seek better explanation for c occurring in i. See Figure 1 for an illustration.

Computing Significance Scores: Consider different students trying to understand section s. We obtain the significance score of a *cphr* c in section t for understanding section s by computing how often these students refer to the description of this *cphr* in section t when reading section s. More precisely, whenever a reader has difficulty understanding a *cphr* in section s and hence is forced to digress to other sections, we compute how likely is the reader to refer to *cphr* c in section t. We then aggregate these likelihoods over many readers and over all *cphrs* in section s. See Figure 2 for an illustration.

We next formalize the algorithmic intuition presented above, and precisely formulate the reader model and the computation of significance scores.

3.2 Notations

Let $S = \{1, 2, \ldots, n\}$ denote the set of sections in a given textbook. Let C denote the set of *cphrs* (concept phrases) in the book. For each *cphr* $c \in C$, denote the set of *cphrs* related to it by $R(c)$. Note that $R(c)$ includes c. Let $\lambda_s(c, t)$ denote the significance score of *cphr* c in a different section t for understanding section s. Let k_s denote the number of desired $\langle$*cphr*, section$\rangle$ concept references in the study navigator for section s. Table 1 summarizes key notations.

3.3 Formulation of Reader Model

We formulate the reader model as a random walk over a concept graph $G = (V, E_p \cup E_d)$. Each node $u = \langle i, c_{ij}, j\rangle \in V$ is

S	Set of sections in the textbook ($\|S\| = n$)
C	Set of *cphrs* (concept phrases) in the textbook
$R(c)$	Set of *cphrs* related to *cphr* c
$\lambda_s(c, t)$	Significance score of *cphr* c occurring in a different section t for understanding section s
k_s	Number of desired concept references to be provided in the study navigator for a given section s

Table 1: Notations

a $\langle$section, *cphr*, position$\rangle$ triplet corresponding to the occurrence of *cphr* c_{ij} in section i and its sequential position j amongst the *cphrs* in the section. Denote the associated section i by $\bar{i}(u)$ and the associated *cphr* c_{ij} by $\bar{c}(u)$. There are two types of directed edges in G. The set of persistence edges E_p consists of directed edges corresponding to sequential reading of the book, that is, there is a directed edge from $\langle i, c_{ij}, j\rangle$ to $\langle i, c_{i(j+1)}, j+1\rangle$ and from the last concept node in a section to the first concept node in the next section. The set of digression edges E_d consists of directed edges corresponding to forced digression, that is, there is an edge from u to v if $\bar{c}(v) \in R(\bar{c}(u))$ (if *cphr* associated with v is related to *cphr* associated with u).

The random walk consists of three types of transitions:

1. Persistence transition: From any node u, follow the persistence edge, that is, the reader persists to read sequentially from the *cphr* occurrence corresponding to u. Denote the probability associated with such a transition as the persistence factor, α.

2. Digression transition: From any node u, follow a digression edge. Denote the total probability associated with a transition along one of the digression edges outgoing from a node as the digression factor, β. Suppose the reader picks a related *cphr* $c' \in R(\bar{c}(u))$. The reader then selects an occurrence of c' amongst all occurrences with equal probability.

3. Return transition: From any node to which the reader has digressed, return to the node from where the digression originated. This transition corresponds to the reader returning back to the starting point after a digression. Denote the probability associated with such a transition as the diligence factor, γ.

The above walk requires keeping track of sequential position of the reader in the book because whenever the reader has digressed, she needs to return to the position from where the digression originated. In other words, the return transition depends not only on the current state in the walk but also the state from which the reader started the digression. The Markov property can be achieved by creating |V| copies of the nodes (and digression edges) as follows. The modified graph consists of the set V of nodes, the set E_p of persistence edges corresponding to sequential reading, and further, a copy of (V, E_d) rooted at each node $u \in V$. The digressions that originate from any node u are confined to the copy of V rooted at u and the return transitions point to u from all nodes in the copy rooted at u. By creating a separate copy of digression edges for each sequential position (node), we implicitly keep track of the state from which the reader started the digression and thus the return transition can be determined based on just the current state.

3.4 Computing Significance Scores

Consider the random digression walk starting from an arbitrary node u (that is, the walk corresponding to the chain of digressions originating from u consisting of only digression and return transitions but no persistence transitions). In this walk, the return transitions always point to u and the digression transitions are determined based on the current state. Hence, this walk induces a Markov chain over the strongly connected component reachable from node u. This Markov chain is (a) finite (b) *irreducible* since the underlying graph consists of a single strongly connected component, and (c) *aperiodic* since the underlying graph is non-bipartite. Thus, the Markov chain satisfies the necessary conditions for applying the fundamental theorem of Markov chains [16], leading to the claim below.

CLAIM 3.1. *There is a unique stationary probability distribution $\pi(u, .)$ associated with the random digression walk starting from any node u in G.*

By definition, the stationary probability $\pi(u, v)$ denotes the probability that the walk starting from node u is at node v in the steady state. In other words, this probability corresponds to the relative frequency with which the reader refers the *cphr* $\bar{c}(v)$ corresponding to v when trying to understand the *cphr* corresponding to u and hence larger $\pi(u, v)$ implies that the reader is more likely to refer to v. Thus $\pi(u, v)$ is a measure of the relative significance of an occurrence of *cphr* $\bar{c}(v)$ in section $\bar{i}(v)$ corresponding to v for understanding the *cphr* corresponding to u. Considering the random walks starting from each concept node in a given section s of the book, we can thus compute the significance of a single occurrence of *cphr* $\bar{c}(v)$ in section $\bar{i}(v)$ for understanding *cphrs* in section s. Our goal is to compute the significance of *all* occurrences of a *cphr* in a section. Hence we further aggregate the above score over all occurrences of *cphr* $\bar{c}(v)$ in section $\bar{i}(v)$. In this manner, we also incorporate the frequency of the *cphr* in the section. Note that we chose not to include persistence transitions for significance score computation since sequential reading is the default reading behavior, and we want to take into account the reader's deviation from this behavior in the form of forced digressions.

We thus define the significance score $\lambda_s(c, t)$ of a *cphr* c in section t for understanding section s in terms of the combined stationary probability associated with nodes corresponding to all occurrences of c in t, summed over random walks starting from all concept nodes in section s. We remark that our definition of $\lambda_s(c, t)$ takes into account the following desired factors: the frequency of c in t, the number of *cphrs* related to c and the likelihood that the description of c in t would be referred for understanding *cphrs* in section s in the book.

DEFINITION 3.2. *Given the stationary probabilities $\pi(.,.)$ associated with the random digression walks, define the significance score of a cphr c in section t for understanding section s as*

$$\lambda_s(c, t) := \sum_{v \in V: \bar{i}(v)=t, \bar{c}(v)=c} \quad \sum_{u \in V: \bar{i}(u)=s} \pi(u, v).$$

In the above definition, the inner summation is over all occurrences of *cphrs* in section s (corresponding to the digressions by readers who are unable to understand different *cphrs* in section s) and the outer summation is over all occurrences of *cphr* c in section t (corresponding to how often these readers refer to the description of c in section t).

3.5 Remarks

Number of concept references: We note that the number of desired references for a section can be determined in multiple ways. It can either be a small fixed number across all sections, or be determined based on the distribution of the significance scores for each section. In the latter case, given a limit k_{max} (say, 5) on the maximum number of references to be shown and a desired coverage κ (say, 75%), we can set k_s to be the minimum of (i) k_{max} and (ii) the number of top $\langle$*cphr*, section$\rangle$ pairs for section s needed to cover κ fraction of the sum of significance scores over all $\langle$*cphr*, section$\rangle$ pairs for this section.

Parameter Values: In our implementation of the reader model, there is effectively one parameter that determines the probabilities of the three types of transitions. When digression originates from a node, there are exactly two choices, to persist reading or to digress, and hence $\alpha + \beta = 1$. Similarly, for subsequent nodes in the digression, there are exactly two choices, to return back to starting node or to digress further, and hence $\gamma + \beta = 1$. Thus $\alpha = \gamma = 1 - \beta$. This relationship between α and γ is in agreement with the following natural intuition: one's tendency to read forward in a section is the same as the tendency to return to the starting point after a digression, since both these tendencies try to achieve the same goal of one's disciplined reading and completion of the entire book.

We experimented with different choices of the digression factor, and confirmed that the results from our reader model are robust to these choices. We chose $\beta = 0.3$ in our implementation. This choice corresponds to the reader starting a digression 30% of the time and persisting to read sequentially 70% of the time.

3.6 Study Navigator with Section References

The study navigator can be generalized to include only section references (that is, references at the granularity of a section) so that each section is treated as an atomic unit of reading. For this purpose, we compute the significance score $\tilde{\lambda}_s(t)$ of section t for understanding section s and then modify our algorithm to return an ordered list of top k section references for section s, based on the significance scores. $\tilde{\lambda}_s(t)$ can be computed either (1) by aggregating the significance scores at $\langle$*cphr*, section$\rangle$ granularity as: $\tilde{\lambda}_s(t) := \sum_{cphr\ c \text{ in section } t} \lambda_s(c, t)$, or (2) modifying the reader model to treat each section as an atomic unit of reading. For example, the reader can be modeled to read an entire section before referring to other sections for *cphrs* that she could not understand. Similarly, whenever she digresses to a different section, she reads the digressed section from beginning to end, and then determines whether to digress to another section or return to the starting section.

Simplified Significance Score Computation: The significance score computation for section references can be approximated using the following simplified algorithm. For each *cphr* c in section s, determine other sections that mention c (say, using the back-of-the-book index if present) and then obtain the significance score of section t for section s as the number of distinct *cphrs* that are present in both s and t. This algorithm tries to simulate a reader who uses the back-of-the-book index to determine other sections to refer to while reading a section. This algorithm uses only information local to a section and other sections that share common *cphrs* while the reader model based algorithm performs a global computation using random walks. The former can be viewed as approximating the latter, analogous to how in-degree (a local measure) can be used to approximate PageRank (a global measure) [12].

4. USER STUDY

We carried out extensive experiments to understand the performance characteristics of the study navigator system and present the results in this section. The goal of our evaluation is to determine whether users find the references provided by the study navigator system useful.

4.1 Methodology

4.1.1 Data Sets

We used a corpus of Indian high school textbooks published by the National Council of Educational Research and Training (NCERT). We selected this corpus because these books were readily available online. The corpus consists of books from grades IX–XII, covering four broad subjects: Sciences, Social Sciences, Commerce, and Mathematics. For the purpose of in-depth analysis, we use Grade XII Economics textbook. We also present results for two other books from very different subjects: Grade X Science and Grade XII History.

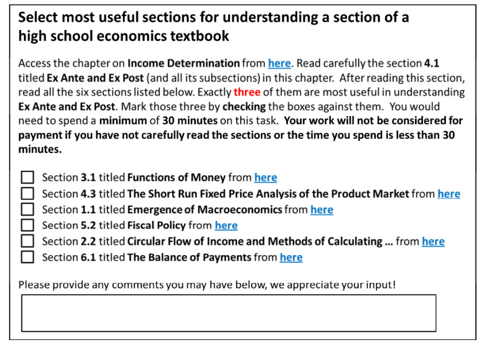
4.1.2 Helpfulness Index

Given the unavailability of a standard benchmark, we used the following procedure to evaluate the usefulness of the references proposed by the study navigator. For a given section, we first determined the top three sections referred by the study navigator. Ideally, we would have liked to compare them with those that an expert human judge (such as a teacher using the book or a student studying from the book) finds most useful after reading the entire book. In the absence of the availability of this subject population to us, we used the Turkers from the Amazon Mechanical Turk platform as judges. However, we could not recruit Turkers who were willing to read the entire book. We, therefore, changed the task to determine if the Turkers can differentiate the sections suggested by the study navigator from other sections. For this purpose, we obtained three arbitrary sections from the book and provided the original section along with these six sections to a judge, after scrambling the ordering between the referred sections. The judge was asked to read the original section, followed by all the six referred sections. Then the judge was asked to select exactly three most useful amongst the referred sections. This exercise was carried out using multiple judges.

We employed Borda's method to merge the votes of different judges. Borda's method strives to achieve a consensus ranking and satisfies desirable properties such as reversal symmetry [21]. Each judge can be viewed as assigning one point each to three out of six referenced sections and zero point each to the remaining three. Denote the total number of points a section obtained from the judges as its vote score. Consider the set of three sections with the largest vote score. These are the sections voted as most relevant by the judges according to Borda's method.

Out of these Borda winners, we determine the number of sections that were also suggested by the study navigator and define the *helpfulness index* as the number of study navigator references in this set divided by three (size of the set). Thus, in the absence of ties, the helpfulness index for each section will be equal to one of the following four values: 1, $2/3$, $1/3$, and 0. A value of 1 means that the top three sections voted by the judges were the same as the top three study navigator section references and a value of 0 means that the judges considered the arbitrary sections as more relevant than the study navigator section references.

However, it may not be possible to uniquely determine the set of three sections with the highest vote because of ties. In this case, we compute the helpfulness index by taking the expectation over all possible choices of this set, as explained in the following example. Let i_1, i_2 and i_3 be the study navigator section references with vote scores of 4, 3 and 3 respectively and r_1, r_2 and r_3 be the arbitrary section references with vote scores of 6, 3 and 2 respectively. The winner set always includes r_1 (section with the largest vote score) and i_1 (section with the second largest vote score). However, there are three possible candidates for the third section: i_2, i_3 or r_2. Thus, possible choices are $\langle r_1, i_1, i_2\rangle$, $\langle r_1, i_1, i_3\rangle$ and



Select most useful sections for understanding a section of a high school economics textbook

Access the chapter on **Income Determination** from here. Read carefully the section **4.1** titled **Ex Ante and Ex Post** (and all its subsections) in this chapter. After reading this section, read all the six sections listed below. Exactly **three** of them are most useful in understanding **Ex Ante and Ex Post**. Mark those three by **checking** the boxes against them. You would need to spend a **minimum** of **30 minutes** on this task. **Your work will not be considered for payment if you have not carefully read the sections or the time you spend is less than 30 minutes.**

- ☐ Section **3.1** titled **Functions of Money** from here
- ☐ Section **4.3** titled **The Short Run Fixed Price Analysis of the Product Market** from here
- ☐ Section **1.1** titled **Emergence of Macroeconomics** from here
- ☐ Section **5.2** titled **Fiscal Policy** from here
- ☐ Section **2.2** titled **Circular Flow of Income and Methods of Calculating …** from here
- ☐ Section **6.1** titled **The Balance of Payments** from here

Please provide any comments you may have below, we appreciate your input!

Figure 3: A sample HIT

$\langle r_1, i_1, r_2\rangle$, with corresponding helpfulness index of $2/3$, $2/3$ and $1/3$ respectively. Hence, the helpfulness index in expectation will be $\frac{2/3+2/3+1/3}{3} = 5/9$. Thus, the helpfulness index for a section can be one of a small set of discrete values. In the results we present, in addition to the expected value, we also provide the two extreme possible values of the index, corresponding to the unfavorable choice (where we favor the inclusion of an arbitrary section in the winner set over a navigator section) and the favorable choice (where we favor a navigator section over an arbitrary section).

4.1.3 Judges

Figure 3 shows the HIT (Human Intelligence Task) provided to the judges. In this example HIT, Sections 2.2, 4.3 and 5.2 are study navigator section references and Sections 1.1, 3.1 and 6.1 are arbitrary sections. Notice that the sections have been randomly ordered.

Each HIT was judged by seven judges. There were 158 distinct judges who took part in the study. We specified that a judge spend a minimum of half an hour on a HIT. We required our judges to have performed at least 1000 HITs in the past with an approval rating of at least 96%. Such judges have a strong interest in retaining their high rating. The judges had at least High School degree. We followed best practices suggested in the literature in accepting HITs [1].

We also validated the quality of judgments along different dimensions. We verified that the judgments did not exhibit position bias and judges were not unduly influenced by the (randomized) order in which the six sections are presented in each HIT. Similarly, we verified that the judges did not have a backward bias (that is, tendency to favor earlier sections in the book) or a forward bias (that is, tendency to favor later sections). See [5] for details.

4.2 Performance Results

The overall performance of the Study Navigator system for the three textbooks is shown in Figure 4. Each book is shown in the X-axis and the helpfulness index, averaged over all sections in the book, is shown in the Y-axis. The extreme values of the index are shown using an error bar. The results are very encouraging. The average helpfulness index for Grade XII Economics and Grade XII History books is 80% and 78% respectively, and this index is as high as 91% for Grade X Science book.

We next show the performance broken down at the section level. Figure 5 gives the fraction of sections with certain helpfulness index for the three books. For 71% of sections in Grade X Science book, the helpfulness index is 100%, that is, the judges considered all three study navigator section references as useful. For over 90% of sections, the helpfulness index exceeds 67%, that is, the judges considered at least two out of the three study navigator section references as useful. For 40% of sections in Grade XII Economics book and for 36% of sections in Grade XII History book, the help-

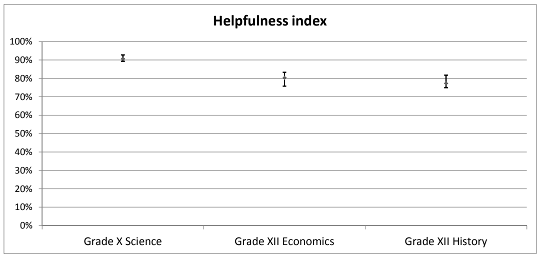


Figure 4: Performance of Study Navigator system for the three textbooks

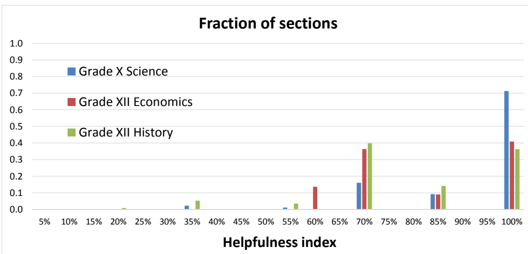


Figure 5: Fraction of sections with different helpfulness index in the three textbooks

fulness index is 100%. Furthermore, for 80% of sections in Grade XII Economics book and for 90% of sections in Grade XII History book, the judges considered at least two out of the three study navigator section references as helpful.

Grade X Science book has a higher helpfulness index because chapters are relatively self-contained in this book. On the other hand, in Grade XII Economics and Grade XII History books, even an arbitrary section can be considered relevant to the original section since common concepts are discussed across many chapters.

We also performed in-depth analysis of the performance results by ourselves reading the books carefully. This analysis revealed that the cases where the judges preferred references other than those provided by the study navigator, it was mostly for sections that reminded them of the material covered earlier, or described applications of the material being discussed, or provided general tools (*e.g.*, how to interpret a graph). See [5] for details.

5. CONCLUDING REMARKS

The future of textbooks is electronic. Sven Birkerts thus opined about this brave new world [7]: "What the writer writes, how he writes and gets edited, printed and sold, and then read – all the old assumptions are under siege." However, the current technology is still quite nascent [23]. We anticipate a surge of innovations to make studying from electronic textbooks much more pleasant and productive. Electronic textbooks as a medium is fundamentally different from printed textbooks, and hence has the potential to enable new kinds of functionalities.

We presented *study navigator*, one such novel functionality that can enhance the experience of studying from electronic textbooks. The goal of the study navigator is to help a student learn the material better and faster by providing easy access to concepts explained elsewhere in the book that are most relevant for understanding the present section.

The study navigator can be adapted to match a student's information processing preference. In an extension of this work in [5], we consider two types of readers: *curious* and *diligent*. When reading a section, a curious student might be open to referring unread later sections that provide advanced information while a diligent student might prefer references only to earlier sections to refresh the material the student has already read. This extension, which we call the *student-specific navigator*, allows students to control the balance between sections that help refresh material already studied vs. sections that provide more advanced information by adjusting a curiosity-factor knob.

Though currently unavailable, rich data on reader's actions can be obtained once electronic textbooks are widely deployed. In the future, we would like to investigate how reader's actions can be incorporated to further enhance the study navigator. Another direction is to explore specializations of the study navigator for additional learning styles beyond the information processing orientations of the reader. More generally, we are interested in designing tools that make use of the electronic format to imbue unique functionality in electronic textbooks along the lines described in the introduction to this paper.

6. REFERENCES

[1] *Amazon Mechanical Turk, Requester Best Practices Guide*. Amazon Web Services, June 2011.

[2] *California Education Technology Task Force Recommendations*. California Department of Education, 2012.

[3] *Report on Aakash tablet*. Indian Ministry of Human Resource Development, 2012.

[4] R. Agrawal, S. Chakraborty, S. Gollapudi, A. Kannan, and K. Kenthapadi. Empowering authors to diagnose comprehension burden in textbooks. In *KDD*, 2012.

[5] R. Agrawal, S. Gollapudi, A. Kannan, and K. Kenthapadi. Study navigator: An algorithmically generated aid for learning from electronic textbooks. Technical Report MSR-TR-2013-68, Microsoft Research, 2013.

[6] R. Bareiss and R. Osgood. Applying AI models to the design of exploratory hypermedia systems. In *ACM Conference on Hypertext*, 1993.

[7] S. Birkerts. *The Gutenberg Elegies: The Fate of Reading in an Electronic Age*. Faber & Faber, 2006.

[8] P. Brusilovsky, J. Eklund, and E. Schwarz. Web-based education for all: A tool for development adaptive courseware. In *WWW*, 1998.

[9] E. H. Chi, L. Hong, M. Gumbrecht, and S. K. Card. ScentHighlights: Highlighting conceptually-related sentences during reading. In *IUI*, 2005.

[10] E. H. Chi, L. Hong, J. Heiser, and S. K. Card. ScentIndex: Conceptually reorganizing subject indexes for reading. In *IEEE Symposium On Visual Analytics Science And Technology*, 2006.

[11] C. Cleary and R. Bareiss. Practical methods for automatically generating typed links. In *ACM Conference on Hypertext*, 1996.

[12] S. Fortunato, M. Boguñá, A. Flammini, and F. Menczer. Approximating pagerank from in-degree. *Algorithms and Models for the Web-Graph, LNCS 4936*, 2008.

[13] N. Henze and W. Nejdl. Adaptation in open corpus hypermedia. *International Journal of Artificial Intelligence in Education*, 12(4), 2001.

[14] D. Jurafsky and J. Martin. *Speech and language processing*. Prentice Hall, 2008.

[15] D. Kelly. Adaptive versus learner control in a multiple intelligence learning environment. *Journal of Educational Multimedia and Hypermedia*, 17(3), 2008.

[16] R. Motwani and P. Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[17] N. Mulvany. *Indexing books*. University of Chicago Press, 2005.

[18] W. J. Ong. *Orality & Literacy: The Technologizing of the Word*. Methuen, 1982.

[19] K. A. Papanikolaou, M. Grigoriadou, H. Kornilakis, and G. D. Magoulas. Personalizing the interaction in a web-based educational hypermedia system: The case of INSPIRE. *User Modeling and User-Adapted Interaction*, 13(3), 2003.

[20] J. R. Remde, L. M. Gomez, and T. K. Landauer. SuperBook: An automatic tool for information exploration – hypertext? In *ACM Conference on Hypertext*, 1987.

[21] D. Saari. *Decisions and elections: Explaining the unexpected*. Cambridge University Press, 2001.

[22] G. Salton, J. Allan, C. Buckley, and A. Singhal. Automatic analysis, theme generation, and summarization of machine-readable texts. *Science*, 264(5164), 1994.

[23] A. Thayer, C. P. Lee, L. H. Hwang, H. Sales, P. Sen, and N. Dalal. The imposition and superimposition of digital reading technology: The academic potential of e-readers. In *CHI*, 2011.

[24] E. Triantafillou, A. Pomportsis, S. Demetriadis, and E. Georgiadou. The value of adaptivity based on cognitive style: an empirical study. *British Journal of Educational Technology*, 35(1), 2004.