A Thousand Papers for ISLIP 97

m.c. schraefel, mc@csr.uvic.ca
Dept. of Computer Science,
University of Victoria

Abstract
The following paper describes the rationale for and implementation design of a
conversational text prototype using Intensional HTML. Currently, IHTML has been used
for dynamic representation of static content. This paper proposes to use IHTML for the
presentation of dynamic conversational content.

The problem:

Texts are usually constructed as static documents. Whether we think of articles published in
journals or on web sites, documents are usually static. Even on the web dynamic requests
still serve up prefabricated, pre-existing whole texts. The problem of what I call Whole
Document Retrieval shows up especially when trying to do an effective search through web
databases. A user does a search engine request: they find a dynamically created list of
documents that may or may not be appropriate. No one document necessarily holds the key
info. Even in the best match document, surplus and redundant info regularly accompanies
the required details.

The problem of search engine query results is also situated as context sensitive searches.
For instance, if I type into Alta Vista "fish food" am I looking for food for fish, fish for
human diet or statistics on discarded employees?

Search engines like Excite and Alta Vista are currently developing approximations of
context searching rather than simple keyword matching. But even were articles to be
retrieved that suite the appropriate context, there is no awareness in those documents of
what I refer to as depth. Depth can refer to level of expertise or amount of information at
what level of expertise. As long as document creation remains static rather than demand
driven, getting information from online sources will continue to be an arduous exercise, no
matter how effective the context matching becomes.

Consider the problem of lack of depth in searching just the Netscape site for information on
how to create frames in web pages, something Netscape invented. this is a fairly narrow
band search. However, even if the search engine retrieves only the "how to build frames"
related documents (excluding all the press releases of same) there is no discrimination for
depth. This means the retrieved articles may be pitched too high for my level of expertise,
in which case I am on my own to find other articles to produce the required info that may
eventually allow me to make sense of the frames article. The converse may also be true: the
retrieved articles may be too basic to answer my more specific questions and again I am on
my own to spend time trying to nail down other sources. That is, there is no way I can
(yet) request the search engine to retrieve the kind of info I want. I cannot request

\textbf{topic: how to build frames}
\textbf{level: expert level}
\textbf{detail: minimum}
I cannot, because document construction is not dynamic. As a reader, I am at the mercy of the text presented and its assumptions about its audience.

Similar examples may be found in course texts or training manuals. If a text for a course or manual is not set at a level for a member of its often enforced audience to understand, the user is handicapped from effective access to the material.

**The Solution: Conversational Texts**

Imagine what would happen if the text could respond to a user's needs so that depths could be set within queries, or within the retrieved documents themselves so that any document can respond real time to the user's queries based on what they happen to be considering in the text at that moment. Imagine a first year calculus student studying derivatives. The text delivering these concepts goes into a discussion of the sin of an angle theta and the student feels lost because they don't remember much about the unit circle from their grade 12 algebra.

In a conversational text, the student could set the depth dimension of the current expertise level down a notch and the depth level of detail up a few notches on the sin concept and generate a gentle review of the unit circle. As confidence with the material increases, expertise and detail can be adjusted accordingly. The instructor (or the text itself) could suggest the expertise level the student should achieve by course end.

In the case of a journal article (like the one you are reading right now), perhaps the only depth level that one would want available is detail where the reader would like to hear more about a particular point. Or less, for that matter. The depth can be raised and detail can be increased.

In these texts, the user determines the interaction with the material. It is this interaction that approximates conversation. The text provides some information; the user wishes to question a point; the text provides a response. Hence, conversational texts.

With the growth of the internet, and therefore the increased presentation of information online, along with the development of intelligent agent searches, such texts could draw on not only local texts for information sources, but eventually, from across the web as well. One could always have "related information" at their fingertips from a variety of voices.

**Conversations among many.**

**Construction and Implications**

Conversational texts allow for many versions of themselves to be constructed dynamically depending on user requests. This means that a variety of audiences need to be constructed for a text as well as a variety of detail. I imagine authors creating info chunks rather than essays for such texts. Depending on the user's request, the appropriate chunks are dynamically assembled which result in the appearance of a single text.

For instance, I might request a text on Frames for an expert user in brief detail. I would receive as output a text based on all the chunks that suited my criteria. I would get one version of all the possible versions of a text that could be created from all the chunks that exist on the topic of Frames creation. All these versions may have certain features in common, like headers and footer information, as well as a working definition of what a frame is, but beyond that, the versions would differ across the dimensions we've described
of expertise and degree of detail. One could possibly create a thousand documents on this
topic depending on the number and level of these dimensions available.

This versioning of texts can be well situated and described within intensional logic
programming and demand driven data flow models. Intensional logic provides a
framework and semantics for describing versions in terms of version space and possible
worlds.

Indeed, based on work by John Plaice and Bill Wadge, the denizens of dinner, it is this
intensional interpretation of the web as demand driven dataflow capable of sustaining
billions and billions of parallel universes (bill's update to possible worlds) that underlines
Taner Yildirim's development of Intensional HTML¹.

One of the strengths of Intensional HTML, unlike standard HTML, is that it assumes and
provides for the necessity of multiple versions of any document entity, from a single web
page to an entire site, or perhaps the entire web. The fundamental tag of IHTML is not the
traditional href, but the #include. The include in IHTML allows web authors to create
document templates that can reference the appropriate text chunks depending on what
versions, based on the combined requested dimensions, have been requested.

Such template creation facilitates the delivery of dynamic texts. The concept of using
IHTML to construct dynamic texts, in particular, dynamic texts that accommodate depth, is
an extension of the original implementation design of IHTML.

In Yildirim's initial presentation of IHTML, one of the amazing features of the
demonstration was the users ability to click on a choice of background colors and have that
choice permeate the site. In traditional HTML, to achieve the same effect, one would have
to clone the entire site and change only the background tag to achieve the same effect².
Similarly, on Yildirim's site, one could change language and have the language change
permeate through the rest of the site. Hence intensional logic's possible worlds become
embodied in IHTML as parallel universes: the site in the French universe is identical to the
English universe except that in the French universe, everything is in French. With the
power of transversional links, one can shift between these parallel universes, constructed
on the fly, whenever one wants.

Such versioning functionality is impressive in itself, making the creation of "text only"
sites, for instance, a far less labor intensive process than formally when all content would
need to be cloned, excising all graphics references. Versioning handles graphics or not as
dimensional settings.

In these implementations, however, the data itself is treated as constant, to be translated
into a different language; with graphics or not; blue background or yellow. Fortunately,
Yildirim's IHTML provides the tools for constructing dynamic texts as described above.
Intensional HTML can accommodate depth dimensions in two ways. In working with
Yildirim for an initial implementation of a conversational text prototype, we evolved the
notion of recursive includes to accommodate one form of depth and, with the input of Bill

¹ Please see “Intensional HTML” by Wadge and Yildirim in this collection.
² Some sites implement Javascript procedures to create this change of color effect. While there are definate
advantages to the browser's interpreting code, in this case, large code chunks must be embedded on every
page that wants to achieve this effect. IHTML's preprocessing eliminates the code load as well as allowing
the change to permeate the rest of the site from that point onward.
Wadge, evolved aggregate inclusion to accommodate chunk selection for depth. Yildirim has implemented both extensions in his latest version of IHTML.

At this point, an example may serve to underline the value of intensional HTML for conversational texts.

A thousand articles on Wuthering Heights

The conversational text site prototype is situated around a chapter from a dissertation on the representation of the relationship between language, the law and who, in a patriarchal culture, can be heard, as demonstrated in the novel Wuthering Heights. The conversational text for this site has the following main dimensions: character, character combinations, issue, global depth.

One can call up a text that presents a detailed analysis(depth) of Cathy (a character) and the Law (an issue). One could call up a quick overview (depth) of Edgar’s relation to Heathcliff (a character combination that is actually a subversion of Edgar, but to the user appears as two character choices) in terms of Language (an issue). For some versions of a document, the material will be very similar to other versions (Cathy and the Law will have much in common with Cathy and Language because of the nature of the material).

The structure may also be similar for viewing these versions. The current structure allows for two main presentations of information. If the user chooses "overview" depth, the site will generate one "page". If the user chooses higher levels of detail, the site’s structure will link to five pages: introduction, main discussion 1, main discussion 2, conclusions, bibliography. The amount of detail in each page will vary, however, depending on depth chosen. Structure of the site could vary entirely, depending on depth chosen, but for my purposes, these two main structures were sufficient.

Textual depth (detail of discussion in this case) is handled by recursively called include files. For this project, it made sense to assume that Depth 10 detail includes the material from Depth 9 which includes material from Depth 8 and so on. However, there may be cases when such a recursively linear approach may not be appropriate and one simply wants to collect and present a collection of all the best finds for the given dimensional attribute request. Such is the case with the bibliography pages, and perhaps represents one of the most immediately valuable offshoots of this project.

When depth is applied to the generation of a bibliography, depth ranges from minimal, which also means most essential to maximal which means broadest selection of material. In these cases, establishing a recursive association among book titles is possible but brittle. It may be true for a list of books that includes only a set number of titles that recursive calls makes sense: the maximal list will defiantly include the members of the minimal list. But what if a book is added or subtracted? How does one insert a new title without screwing up the order of the books called? In this case, aggregation is used to pull together all the titles marked to fit the particular criteria and a small program is then called to order the books by author in either ascending or descending order.

As the number of participants contributing chunks over time to a project evolves, aggregation and ordering may entirely replace recursive includes since an increasing number of alternate text chunks would create the same problems for presenting as the changes to titles in a biogeography. For now, however, recursion is efficient where there is no intention to add any more to the text beyond its current chunks.
The immediate problem that recursion solves over aggregation for this kind of text serving is order: for a text to make sense, each chunk must usually form some type of relation to its neighbor. How is this ordering to be established among a group of chunks of the same depth level where an alphabetical or numerical cue is not readily available?

One possible solution of course is to leverage the effect writing chunks actually has on the writers conception of the given project. Chunks become aphorisms of one or several paragraphs. ordering of the aphorisms within a related group becomes largely incidental. While this may sound somewhat mystical, this form of writing was consistently practiced by a very unymystical philosopher, Nietzsche. Nietzsche's writing consist of just these types of aphorisms. In the appendix of my master's thesis I pulled all Nietzsche's aphorisms I could find from all his works on the subject of Women. The result was a new Nietzsche text, Nietzsche on Women. The read was fascinating when so assembled. The effect is cohesive, but it was only possible because Nietzsche wrote in a way to facilitate just such versioning.

With such writing at its heart, intensional texts could be constructed from the aggregation of aphorisms.

The advantages of conversational texts over traditional static documents and the problems of bloated Whole Document Retrieval are many. On the readership/user side, they allow one source site to effectively serve many audiences. On the text creation side, they can readily facilitate collaborative authorship where various authors can be assigned a set of chunks to construct that may be called together in ways they did not anticipate to create new texts and new knowledge.

In many respects the above description suggests approaching writing projects like object oriented software: constructing components that can be reused, that can inherit, that can be extended for a variety of instances. This is very true, not the least because the engine that drives these intensional, versional demands, HTML, is fundamentally oo'ish in its adoption of inheritance. But another metaphor that may serve derives from Dale Spender's consideration of what the new media online writing implies for authorship. She sees writing becoming increasingly like film making. In a single film, there are many many contributors to the final product. Where powerful software allows more people access to digital effects tools, which can reduce the number of persons required for any project, the same tools facilitate the development of collaborative projects. In a world of intensional, conversational texts, I may be able to contribute valuable chunks to a variety of projects rather than develop one single, monolithic tome. In one project, I may act as chunk editor, on another, graphic chunk coordinator or template/structure designer.

As stated, the benefits of truly dynamic documents are as many as their are readers. Intensional modelling that facilitates depth may become the best approach for their construction.

---