A Commentary on Frank Halasz's "Reflections on NoteCards: Seven Issues for the Next Generation of Hypertext Systems"

Elli Mylonas
Associate Director for Projects and Research
Scholarly Technology Group
Brown University
Providence, RI 02912
elli_mylonas@brown.edu

Abstract

Although hypertext research and development didn't continue exactly in the directions that Frank Halasz predicted in 1988 and 1991, much of his analysis and agenda is still valid. The most interesting development is the shift away from hypertext systems as applications software to either embedded functionality or information structuring. The focus on standards and protocols has led to progress on Halasz's issues and has made hypertext almost universal. On the other hand, the ubiquity of hypertext on the Web is a force that works against some of Halasz's agenda. It is difficult to tailor or extend a hypertext because the task requires knowledge of the tools, the information structures, and the content.

Keywords

hypertext/hypermedia, World-Wide Web, systems, standards, protocols, structures, content

Introduction

Frank Halasz's "Reflections on NoteCards" (Halasz, 1988), fondly known as the "Seven issues" paper, is an early article that sets an agenda for hypertext research and development. It may be better known within the hypertext community in its incarnation as an early member of a series of thought-provoking *tours de force* that have come to characterize the hypertext conference keynote addresses. In setting forth his agenda, Halasz had to analyze and categorize the state of the art of hypertext at about 1987. Such an article, if it dwells too much on technical aspects of the agenda, can easily become dated. From its early days, however, hypertext research has had as its goal to enhance and facilitate intellectual work by means of the technology it developed. Halasz does the same. There are two ways to consider how the article has stood the test of time. One is to see whether his suggestions for future research

ACM Journal of Computer Documentation, 2001,25:104-108
 ISSN 1527-6805

and development happened. The other is to see if the same criteria and directions can be applied to today's hypertexts. I will selectively try both approaches.

As I reread Halasz's article, I also debated briefly whether I should consider it in terms of the type of hypertext systems that Halasz was citing in 1988, or if I should try to tease out the aspects of the World-Wide Web that the article can be measured against. I decided to go with the Web because it is a "system" which has more users, in all its forms, than any other. More importantly, it has become the universal interface. Many research projects use the Web as their distribution system and interface. Also, Web research is one of the main areas presented at the series of Hypertext conferences, where hypermedia research has traditionally been published, and more and more traditional hypermedia research is presented and cited at the Web conferences. It seems clear that the Web and the hypermedia sphere intersect and are related.

The "Seven issues" paper already comes with its own critique; Halasz revisited and revised his own seven issues in the closing keynote address at Hypertext 1991 (Halasz, 1991). In the later paper, he rejects and redefines his previous issues together with the dimensions against which a hypertext system should be categorized. Halasz also points out that the article does not express a 1988 viewpoint, but rather is a revision of his 1987 Hypertext conference paper (Halasz, 1987), which was begun in 1986 (Halasz, 1991, slide 1-2). The 1991 keynote address refers to parts of the 1987 address and the slides that accompanied it. These three papers form an interrelated set and cannot be read independent of one another. In this discussion, I will be referring to all three.

In order to discuss the seven (or later, ten) issues, it is helpful to study the definition of hypertext in the paper and see how it applies to current hypertext systems. Halasz originally defines hypermedia as "a style of building systems for information *representation* and *management* around a network of multi-media nodes connected together by typed links" (Halasz, 1988, p.836). In the keynote address, he realizes that the definition is too narrow and redefines it as follows:

Hypermedia is a style of building systems for the creation, manipulation, presentation and representation of information in which:

- the information is stored in a collection of multimedia nodes
- the nodes are explicitly or implicitly organized into one or more structures (commonly, a network of nodes connected by links)
- users can access information by navigating over or through the available information structures (Halasz, 1991, slide 2-4).

The first definition assumes that hypermedia is embodied in a system, and it reflects the view that it will be used to actively manipulate information. The second definition is more general; it allows for presentation of information as well, and easily applies to contemporary hypertext systems, even to most systems we use today that contain hypermedia capabilities. It still assumes the presence of a system, but it practically dispenses with the notion of the link altogether. The term *node* is left vague, so at first glance it appears that hypertext has to be chunk style, but later in the paper Halasz discusses hypertexts consisting of longer documents that you can "fly" through (Halasz, 1991, slide 2-11). So the limited node of NoteCards or Intermedia is not a requirement.

The part of both definitions that has not stood the test of time is the concept of the hypertext or hypermedia system. Hypertextual features are an expected part of many applications, but applications that can be characterized primarily as hypermedia applications are rare. The most widespread type of hypertext today, the Web, is not a system, and it isn't even always hypertext. It is best characterized as an environment, and hypertext as a type of expression and interaction. Open hypermedia systems (OHS), the most system-oriented domain of current hypertext development and research, is related to the second of the 1991 revised issues: open systems (slide 4-5 ff.). Halasz's intent, which was the same as that of the creators of Microcosm (Davis et al., 1992), Devise Hypermedia framework (Gronbaek and Trigg, 1996), and others, was to find a way to free users from the isolating, monolithic hypertext system. Open systems served to decouple the linking service from the information representation and creation. In 1988, the only way one could use hypertext was by building a hypertext system. Today, the focus of the hypertext domain is primarily on the design and manipulation of information. Significantly, this is also one of the concluding suggestions of the keynote address.

Selected Seven Issues over Time

In the 10 years since Frank Halasz's article and his keynote address, hypertext systems as a topic of

research waxed and waned (not always in that order) and developed in different directions. A number of research areas that are part of the seven issues became prominent and even characteristic of the field. For example, the ability to search and query (Halasz, 1991, issue 1) in a hypertext system is taken for granted. Supplemented with the ability to move around a hypertext using tools other than navigation, links also are not a required and defining part of hypertext (Halasz, 1991, issue 1).

Some issues were never developed, or they moved into other research domains than hypertext. An example of this is collaborative work, which although still deeply related to hypertext research, is studied in its own domain: computer-supported collaborative work. Very large hypertexts also are no longer an active part of hypertext research, although they do exist. They are either taken for granted or handled over in a very large database (VLDB). Halasz himself removed versioning from the seven issues altogether in 1991, because little work had been done on the topic and it no longer seemed important. Still, recent work on WebDAV, the standard for collaborative authoring on the Web, is finally leading to the creation of hypertext systems (Web servers) that have integrated versioning capabilities, even if they are much simpler than the ones envisioned by hypertext researchers.

The rhetoric of hypermedia, which Halasz noted as an issue worthy of study in 1991, has become a very important research area in its own right. Literary hypertext and hyperfiction have become important as they use hypertext and try to understand how to read and write in this medium. Papers on these topics are presented not only at scientific hypertext conferences, but at disciplinary conferences in the humanities. A rich cross-fertilization has developed from the seeds planted by George Landow, Stuart Moulthrop, Michael Joyce, and others (Landow, 1987; Moulthrop, 1992; Bernstein et al., 1992), where the "Literati" and the "Engineers" look at each other's work to learn more about their own. The Hyperfiction reading event is common and well attended at hypertext conferences.

In the 1990s there were many research hypertext systems, but few of them became commercial, and none were widely adopted. Exceptions to this rule occurred within specific domains, where, for example, Eastgate Systems' StorySpace was widely used by students in English composition courses and by writers of hypertext fiction. Hypertext started to become a generally available facility at the same time. Help systems on desktop computers were rigidly hierarchical hypertext systems, for example, and word processors provided the ability to create and follow links. The growth of multimedia applications, especially for education, also relied on hypermedia linking and navigation.

In his last slide in 1987, Halasz said that a measure of the success of hypertext would be if it ceased to be an issue, because it had become so commonplace. In 1991, he decided that it wasn't going to disappear and still needed a research agenda. However, the question of whether hypertext has disappeared is still a good one. It is ubiquitous, because the digital medium allows easy connections and navigation between documents of different types. However, it is also possible to consider the Web, ubiquitous and under constant development, as the biggest and most popular open system yet. Coincidentally, an early presentation of the infant Web was as a demo at the Hypertext '91 conference.

The World-Wide Web

The World-Wide Web, like other systems that Halasz describes such as emacs and Hypercard, does not inherently lead to the creation of hypertextual documents and document collections. It provides a framework and toolkit with which many kinds of systems may be constructed. This encourages issue 7, extensibility and tailorability (Halasz, 1988, p. 850). Often, the Web is used as a general presentation and authoring system, and need have no hypertextual features at all, or none other than those available in conventional printed material. This discussion should only include the hypertextual aspects of the Web represented by hypertext publications and research systems that use it as a framework.

Hypertext as embodied by the Web also lies at some extremes of the hypertext dimensions such as browsing vs. authoring (Halasz, 1988, p. 840). Some of the other dimensions—task specificity, and from the keynote address, Navigators vs. Architects, Card Sharks vs. Holy Scrollers, and Literati vs. Engineers—are still very handy in categorizing different constructs on the Web. Like many early hypertext systems, the Web was developed for an application domain, in order to enhance the creation and sharing of ideas via a collaborative, annotative system for sharing publications. The difference is that the Web is not a system or a particular

piece of software, but a series of standards and protocols, which is where its power resides. One of the defining features of the first presentation of the Web was its use of SGML. Since that time, the World-Wide Web consortium has become one of the important movers in standards development for information structuring and manipulation in the world.

In 1991, the most significant extension to the seven issues is the addition of three "marketing" issues. These have to do with the penetration and acceptance of hypertext systems and documents. The widespread use of the Web is in the process of defining the hypertext market. It seems clear that there is a market for hypertext, although its economics are unclear. The tenth point, publishing hypertexts, is thriving, although not only in the traditional publishing economy. Mark Bernstein, in his Hypertext '99 keynote address, demonstrated that many hypertexts are being published (Bernstein, 1999). They are self-published and often not literary, but they are being written, published, and read..

Original Seven Issues	Revised Ten Issues
1. Search and query 2. Composite nodes 3. Virtual structures for dealing with changing information 4. Computation in/over hypermedia networks 5. Versioning 6. Collaborative work including rhetoric of hypermedia 7. Extensibility and tailorability	Technology: 1. Ending the tyranny of the link 2. Open systems 3. Support for collaborative work 4. User interfaces for large information spaces 5. Very large hypertexts 6. Tailorability and extensibility 7. Computation in (over) hypermedia networks Marketing: 1. Defining the hypermedia market(s) 2. Standards 3. Publishing hypertexts

Primacy of Information

Information and standards have moved into a central position in the Web. The term *hypertext*, when Halasz was writing, referred to a system, and the group of people who knew much about it was small. Today it's universally recognized, and refers to a means of expression or publication. (It's also incorrectly but widely used as a synonym for HTML codes.) So the locus of hypertext has shifted from the system to the contents, which reflects a shift from emphasizing the work of the engineers to the work of the content creators. It lets information choose how it will be viewed and doesn't hold it prisoner to any one system.

The Web's reliance on standards and protocols has created an environment in which the capabilities of a hypertext system are separated from the problems of either creating the information or presenting it. Several of the original seven issues are being addressed, although in a simple way at the outset. I should make it clear here that I am not discussing the Web in its raw form, using HTML 4.0 and standard browsers, but the ability to have servers with hypertext functionality manipulating and distributing content that is richly coded in XML. In this scenario, it is possible to search and query not only content but also structure. The Dynaweb software originally developed by Inso Corporation already does this, and the XPath and XPointer working groups are developing ways to identify and point to parts of documents with linking primitives. Alternatively, search engines like Google use link structure as well as content to find relevant answers to queries (Brin and Page, 1998). It is also possible to cre-

ate and manipulate composite nodes, by defining XML constructs for them. Virtual structures are also more easily arrived at because of the decoupling of the data from the delivery system.

Conclusion

Although hypertext research and development didn't continue exactly in the directions that Halasz predicted in 1988 and 1991, much of his analysis and agenda is still valid. The hypertext community, the overlapping Web research community, and even the domain of digital libraries could use the seven issues to plan an agenda and the seven dimensions to measure their work. The most interesting development is the shift away from hypertext systems as applications software to either embedded functionality or information structuring. The focus on standards and protocols has become the way to make progress not only on Halasz's issues, and in the process has made the use of hypertext almost universal. On the other hand, the very ubiquity of hypertext as represented on the Web is a force that works against the pursuit of some of Halasz's agenda. It is still difficult to tailor or extend a hypertext, not just because the tools require special knowledge, but because it is necessary to have a thorough knowledge of the information structures and content before working with it.

Halasz's article taken together with his 1991 keynote address has survived the advent of a universal hypermedia system and continues to provide a fruitful framework for evaluating and developing hypermedia. It would be very interesting to revisit it again in 10 more years.

References

Bernstein, M., Joyce, M., and Levine, D. (1992). Contours of constructive hypertexts. *Proceedings of the ACM Conference on Hypertext, 1992.*

Bernstein, M. (1999). Where are the hypertexts? Asking again (opening keynote address). ACM Hypertext '99 Conference. [http://www.eastgate.com/ht99/slides/Welcome.htm]

Brin, S., and Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Proceedings of the 7th International World Wide Web Conference*, April 1998.

Halasz, F. (1987). Reflections on NoteCards: Seven issues for the next generation of hypermedia studies. *Hypertext '87 Papers.*

Halasz, F. (1988). Reflections on NoteCards: Seven issues for the next generation of hypermedia studies. *Communications of the ACM,* 31(12), 836-852.

Halasz, F. (1991). "Seven issues": Revisited (closing keynote address). ACM Hypertext'91 Conference. San Antonio, TX. [http://www.parc.xerox.com/spl/projects/halasz-keynote/]

Davis, H., Hall, W., Heath, I., Hill, G., and Wilkins, R. (1992). Towards an integrated information environment with open hypermedia systems. *Proceedings of the ACM Conference on Hypertext, 1992.*

Gronbaek, K., and Trigg, R. (1996). Toward a Dexter-based model for open hypermedia: Unifying embedded references and link objects. *Proceedings of the ACM Conference on Hypertext, 1996.*

Landow, G. P. (1987). Relationally encoded links and the rhetoric of hypertext. *Proceeding of the ACM Conference on Hypertext, 1987.*

Moulthrop, S. (1992). Toward a rhetoric of informating texts. *Proceedings of the ACM Conference on Hypertext, 1992*, pp. 171-180.

StorySpace, Eastgate Systems. [http://www.eastgate.com]

World-wide web distributed authoring and versioning (WebDAV). Internet engineering task force (IETF) Request for comments (RFC) 2518 [http://www.ics.uci.edu/~ejw/authoring/]

XPath [http://www.w3.org/TR/xpath]

XPointer [http://www.w3.org/TR/xpointer]