Informatiewetenschap 1997 : wetenschappelijke bijdragen aan de 5e Interdisciplinaire Conferentie Informatiewetenschap, Eindhoven, November 27, 1997

Eindhoven University of Technology
Department of Mathematics and Computing Science

Werkgemeenschap Informatiewetenschap

INFORMATIEWETENSCHAP 1997

Wetenschappelijke bijdragen aan de
Vijfde Interdisciplinaire Conferentie Informatiewetenschap

Redactie:
Prof.dr. P.M.E. De Bra

97/18

ISSN 0926-4515

Reports are available at:
http://www.win.tue.nl/win/cs

Computing Science Reports 97/18
Eindhoven, December 1997

Werkgemeenschap
Informatiewetenschap

INFORMATIEWETENSCHAP 1997

Wetenschappelijke bijdragen
aan de
Vijfde Interdisciplinaire
Conferentie Informatiewetenschap

Redactie:
Prof.dr. P.M.E. De Bra

Eindhoven, 27 november 1997

INHOUD

Voorwoord

De Interdisciplinaire Conferentie Informatiewetenschap is in 1997 aan haar vijfde editie toe. Ze brengt (vooral Belgische en Nederlandse) onderzoekers, deskundigen, probleemeigenaren en andere geïnteresseerden op het vakgebied "Informatiewetenschap" bij elkaar.

De eerdere conferenties in deze serie werden gehouden in Nijmegen (1991), Enschede (1992), Tilburg (1994) en Delft (!996). Deze conferenties zijn aanvankelijk gehouden onder auspiciën van StinfoN, de Stichting Informatiewetenschap Nederland. Later heeft de (door StinfoN opgerichte) vereniging "Werkgemeenschap Informatiewetenschap" de organisatie van de conferentie overgenomen. In 1996 is besloten om de frequentie van de conferentie te verhogen, teneinde beter in te kunnen spelen op de steeds sneller wordende evolutie van dit wetenschapsgebied. De conferentie Informatiewetenschap'97 is de eerste uit de "oneven" reeks, waarin Belgische en Nederlandse onderzoekers en ontwikkelaars een overzicht geven van onderwerpen die ze bestuderen.

Een conferentie als deze kan slechts een succes worden dank zij de medewerking van velen. Ik spreek gaarne mijn dank uit aan de auteurs van de ingezonden bijdragen en de beoordelaars van de Programmacommissie. Een bijzonder dankwoord richt ik tot mevrouw Tonja van Hoek, die de gehele administratie van deze conferentie heeft verzorgd, alsook deze bundel van conferentiebijdragen.

Ik wens alle deelnemers een vruchtbare conferentie.

Paul De Bra
Voorzitter conferentie Informatiewetenschap 1997

Programmacommissie

Paul De Bra* (Technische Universiteit Eindhoven)
Egbert De Smet (Universitaire Instelling Antwerpen)
Lynda Hardman (Centrum voor Wiskunde en Informatica)
Kees van der Meer** (Technische Universiteit Delft)
Paul Nieuwenhuizen** (Vrije Universiteit Brussel)

* Ook deeltijds Universitaire Instelling Antwerpen en Centrum voor Wiskunde en Informatica.
** Ook deeltijds Universitaire Instelling Antwerpen.

Linguistic Variation in Information Retrieval and Filtering

A. Arampatzis*, C.H.A. Koster, P. van Bommel, Th.P. van der Weide

Technical Report CSI-R9701, January 1997

1 Introduction

The tremendous increase of networked information has led to a new challenge in "information seeking". Currently, users everyday confront themselves with large amounts of information in the form of news, e-mail messages, and especially World-Wide Web pages. Although users of this electronic information have access to a rich body of information, only a small fraction of this is actually relevant to the interest of any particular user. In order to reduce the effort of a user determining which information is relevant to his needs, an automatic solution seems indispensable. Assuming specific long-term interests of a user, and taking into account that the dynamic and unstructured information sources have a high modification rate, this information filtering problem differs from the classical information retrieval problem ([BC92]). However, many of the techniques used for information retrieval can easily apply to filtering and vice versa.

Due to the fact that the largest amount of this information consists of text documents, many approaches have been seen in text filtering which all have in common four basic components:

- a technique for representing documents
- a technique for representing the information need (*profile or query*)
- a way of comparing profiles/queries to document representations
- ways of using the results of comparison (rendering, presentation, interaction and feedback)

Figure 1 illustrates the representation and comparison process implemented by text filtering systems. A framework for the text filtering problem can be found in [OM96].

The state-of-the-art text filtering/retrieval systems are mostly based on use of keywords, both in representing information objects and as a basis for the retrieval language (expressing the information need). The possibility for further improvement in Precision and Recall based on keywords is rather marginal. Besides, the use of keywords is inadequate for more inflected language than English. Citing C.J. van Rijsbergen [Rij79]:

*Dept. of Information Systems, Faculty of Mathematics and Computing Science, University of Nijmegen, Toernooiveld, NL-6525 ED Nijmegen, The Netherlands, E-mail: avgerino@cs.kun.nl

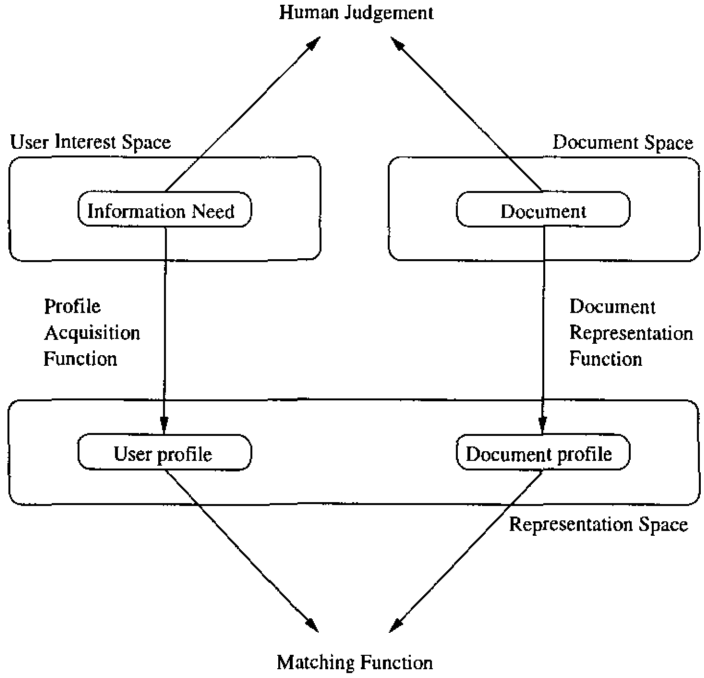


Figure 1: Text filtering model

> *"A big question, that has not yet received much attention, concerns the extent to which retrieval effectiveness is limited by the type of document description used. The use of keywords to describe documents has affected the way in which the design of an automatic classification system has been approached. It is possible that in the future, documents will be represented inside a computer entirely differently."*

In the information filtering project PROFILE ([HSB+96]) linguistic techniques will be used for characterizing documents and for formulating user profiles and queries. For other filtering approaches based on natural language see e.g. [Ram91] and [Ram92]. The approach described in [SO95] also comprises a natural language processor. Our approach is based on the use of *noun phrases* (NP's) instead of keywords. For previous work related to the use of NP's in information retrieval, refer to e.g. [AT96], [ATK96]. The same linguistic techniques incorporated in PROFILE can easily be adapted to other information retrieval and filtering systems.

2 The Information Filtering Project PROFILE

2.1 Organizational Structure

The information filtering project PROFILE is conducted in cooperation between two research groups of the Nijmegen University:

- The Software Engineering group of the Computer Science Institute (CSI), and
- The Cognitive Ergonomics group of the Nijmegen Institute for Cognition and Information (NICI).

The collaboration of these groups may be visualized as in figure 2. The figure splits the Profile

NICI		CSI
Modeling	$\Rightarrow$ *profile* $\Rightarrow$	**Parsing**
$\Uparrow$ *user behavior* $\Uparrow$	Filtering project PROFILE	$\Downarrow$ *descriptors* $\Downarrow$
Interaction	$\Leftarrow$ *documents* $\Leftarrow$	**Retrieval**

Figure 2: Organizational and Functional structure

project into four sub-projects and also delimits the different areas of research.

2.2 Functional Structure

Goals and interests of an individual user or a group of users are used to build a *profile.* Profiles are utilized to support users in formulating queries, to better understand the meaning of them, and for comparison with documents in order to filter out irrelevant information. The *User Modeling* module is responsible for providing these profiles, by building a mechanism to infer information needs from goals and interests, as well as a generator to translate information needs into vectors of noun phrases. This information need is described using natural language phrases, since users are able to verbalize easily their need.

In order to extract useful information, profiles have to be parsed. The *Parsing* module transduces noun phrases to *phrase frames* (also called *information descriptors*) which are used for retrieval.

The *Retrieval* module ([WBHW97]) employs *autonomous intelligent information agents* to collect documents from information sources, which are also reduced to information descriptors. The matching process selects documents to be presented to a user.

The *User Interaction and Rendering* module displays the information in the right way, logging users' reactions about presented information. This can implicitly and/or explicitly give relevance feedback to the user modeling for updating and refining profiles, adjusting the filters better to users' need. The interaction module provides the interface between users and the system, whenever this is needed, in a user-friendly and easy-to-understand way.

The Rest of the Article

The rest of this article in organized as follows. In section 3 the Parsing Engine of PROFILE is described. There is a discussion about how noun phrases are used for retrieval, how linguistic variation is taken into account, and about the notion of information descriptors. In section 4 the research questions and approach are stated. These are the properties which an IR/IF parser must possess, and the normalization and similarity issues of noun phrases. The conclusions are drawn in section 5.

References

[AT96] A.T. Arampatzis and T. Tsoris. A Linguistic Approach to Information Retrieval. Master's thesis, Department of Computer Engineering and Informatics, University of Patras, Patras, Greece, June 1996. Available from: http:// www.cs.kun.nl/ ~avgerino/ LA2IR.ps.Z.

[ATK96] A.T. Arampatzis, T. Tsoris, and C.H.A. Koster. IRENA: Information Retrieval Engine based on Natural language Analysis. Technical report CSI-R9623, Computing Science Institute, University of Nijmegen, Nijmegen, The Netherlands, 1996.

[BC92] N.J. Belkin and W.B. Croft. Information filtering and information retrieval: Two sides of the same coin? *Communications of the ACM*, 35(12):29–38, December 1992.

[HSB+96] E. Hoenkamp, L. Schomaker, P. van Bommel, C.H.A. Koster, and Th.P. van der Weide. Profile - A Proactive Information Filter. Technical Note CSI-N9602, Computing Science Institute, University of Nijmegen, Nijmegen, The Netherlands, 1996.

[OM96] D.W. Oard and G. Marchionini. A Conceptual Framework for Text Filtering. http:// www.ee.umd.edu/ medlab/ filter/ papers/ filter.ps, 1996.

[Ram91] Ashwin Ram. Interest-based information filtering and extraction in natural language understanding systems. In *Bellcore Workshop on High-Performance Information Filtering*, Morristown, NJ, 1991.

[Ram92] Ashwin Ram. Natural language understanding for information-filtering systems. *Communications of the ACM*, 35(12):80–81, December 1992.

[Rij79] C.J. van Rijsbergen. *Information Retrieval.* Butterworths, London, United Kingdom, 2nd edition, 1979.

[SO95] H. Sorensen and A. O'Riordan. A learning personalised information filter. In *Proceedings of the AI'95 Conference*, Montpellier, France, 1995.

[WBHW97] B.C.M. Wondergem, P. van Bommel, T.W.C. Huibers, and Th. van der Weide. Towards an Agent-Based Retrieval Engine. In J. Furner and D.J. Harper, editors, *Proceedings of the 19th BCS-IRSG Colloquium on IR research*, Aberdeen, Scotland, April 1997.

Task-Based Information Filtering: Providing Information that is Right for the Job

Paul De Bra, Geert-Jan Houben, Frank Dignum
Department of Computing Science
Eindhoven University of Technology
{debra,houben,dignum}@win.tue.nl

Abstract: Many attempts have been made to provide Internet and Intranet users with tools that aid them in finding valuable information in the many gigabytes of data they have access to. And although large search engines like Alta Vista and Excite sometimes find the appropriate documents, based on just a few well-chosen keywords, most of their answers are not relevant for the user.

Many company Web-servers are beginning to offer search engines right on the first page, to guide visitors to the information they are looking for. Although the overload of irrelevant information from these services is less than with the global search engines, it is still difficult to find the information a user wants.

The core of the problem with these search engines is their *one size fits all* approach to information retrieval. We propose a different strategy: by using an agent architecture that distinguishes three types of agents (process agents, document warehouse agents and retrieval agents) we can take into account the role of the user in her organization, or the task for which she needs the information. In order to evaluate which documents are relevant for which tasks we propose that cooperative retrieval agents learn to select appropriate documents based on user-feedback.

1. Introduction

Every World Wide Web user has experienced the problem of finding relevant information. Neither subject-based menu systems like Yahoo, nor large search engines like Alta Vista and Excite provide a way to quickly find the documents a user is looking for. Even when one locates a valuable site it is often still difficult to find the appropriate documents on that site. Many Web-servers try to overcome this problem by providing their own miniature version of the large search engines. The information overload on a single site is less dramatic of course, but finding the right documents can still be a problem even on a single site.

The core of this problem is that all available search tools select documents based on the *textual content* of the document, and not on the *purpose* or task the document is written for. When one connects to a typical Web-server, information is usually presented based on the hierarchical structure of the company or organization. For most visitors this structure is irrelevant. A presentation based on *who* the users are or what the *purpose* of their visit is would greatly help most users. But there is still a danger that none of the offered choices matches the reason why a user contacts the site.

We lack a good mechanism to manage an organization's information in such a way that users have easy and efficient access to the information that is relevant for their tasks. This information (management) system should support three aspects of usage:

- helping the users in their access to information: finding the information
- helping the user community to manage and maintain the information: organizing the information warehouse
- helping the user community to replenish the information: updating or adding new information

In this paper we concentrate on the first of these aspects: supporting the users in finding information. It is essential to acknowledge the relationship between

- the place of an activity within a business work process, and
- the need for information during the execution of the activity.

In [HD97] we have described how agents can be used to support the work processes and their activities. These agents contain knowledge about the goals of the process and the standard procedure to fulfill that goal. They also contain knowledge about which information is needed for each step in this standard procedure. Besides the knowledge about the standard procedure they contain a planning module that can be used to construct a plan to reach the goal of the process in those cases when the standard procedure cannot be followed. Here we combine these agents with agents that support the users in finding and receiving information. Thus we construct an information system that supports the enterprise-wide exchange of information.

Specifically, we propose the following cooperation between the process agents and the retrieval agents. When the process agent needs information to support the next step in a business-process it will not only send this request to the retrieval agent, but will also provide information about the context of this request. I.e. it will indicate the goal of the process and the role of the information in the activity to reach that goal. In this way the retrieval agents can build up a user-profile not only based on the word-usage of retrieved documents, but also based on the context in which the documents are used by the user. Thus our agents learn *why* certain documents are considered relevant by the user.

2. Task-Based Information Retrieval

In an environment like World Wide Web, but also in enterprise-wide information systems (e.g. Intranet solutions) in any medium to large sized organization, information is available on a wide variety of topics. The information comes from many different sources and is used by very different kinds of people. Both the menu-based systems like Yahoo and the huge search engines like Alta Vista and Excite are purely *subject* oriented. They try to meet the challenge of providing pointers to valuable documents, based on a search pattern which often consists of just a few keywords.

Many approaches exist to improve on this kind of search technique, by using information from more than just a single user query. Golovchinsky [G97a,G97b] assigns weights to search terms based on how many queries ago the search term was used. Queries in his system are actually hypertext links, not user-typed sets of keywords. Fishnet [BL97] is a tool, developed at the Eindhoven University of Technology, that is typical for agent-based retrieval tools that maintain a database of representations of previously returned accepted and rejected documents, in order to form a user model that represents the typical *interest* of the user. All these types of tools classify documents based on *content*.

We argue that the above mentioned tools cannot provide satisfactory query results because whether a document is relevant or not cannot be easily determined based on a document's content. When a user asks for ''automobile repair'' a search engine will return documents with hobby repair instructions for various engine problems, detailed instructions for experienced car mechanics, help information on auto-body work, addresses of repairmen and shops, etc. Whether documents are relevant to the user depends on much more than just the subject of the document:

- *Who* is the user, what is her job, her training, her skills?
- Which *task* is the user trying to accomplish when asking the query?
- *Where* is the user located (and/or where is she trying to go)?
- At *which company* (or organization) is the user?

All these aspects are related to the specific role the user is playing within the work process. A factory organization supplies its shop floor workers with the proper material (parts, tools, etc.) based on the position of the workers in the production process. (E.g. the carpenter and the designer get different pencils.) In the same way an administrative organization must also supply its workers with the proper material (information, documents, etc.) that is suited for their role in the administrative process.

In order to realize this we propose retrieval agents that use:

- knowledge about (the state of) the process
- feedback from the users about the relevance of the supplied documents

For the process knowledge the retrieval agents should communicate with the process supporting agents (see [HD97]). These process agents use an approach like Action Workflow [MWFF] to establish knowledge about the state of the process. The retrieval agents should learn from the process agents about the state of the process, and therefore about the (business) purpose of supplying the information.

The retrieval agents should ask for user-feedback in order to learn what characterizes relevant documents and irrelevant ones. The difference with others is that in our proposal the user-feedback is not limited to a boolean "relevant/not relevant" selection, but it includes feedback on:

- *topic*: does this document deal with the requested subject?
- *job/level*: is this document appropriate for the user's job and is the material at the right (skill) level?
- *task*: is this document helpful for the user's task?
- *location*: is this document useful for users in this (geographic) area?
- *organization*: is this document useful for users in this company/organization?

Some of this feedback can be given automatically by the process agent, while other aspects should be asked from the user or learned through experience. By discriminating documents according to these different criteria, a search agent is not only able to better find information that is actually helpful for the user, in her job situation, but agents can be tied together to form a more detailed classification of information, as is described in the next section.

Moreover, this structured approach gives a better tool to organize the information and document management process. Any standard factory organization invests in setting up the right support mechanism to facilitate the supply of material to its workers; the average administrative organization on the other hand does not properly acknowledge the different activities that should be involved in supplying the workers with the right documents. For example, document warehouse management is not something that can be completely left to automated agents (just as there are only exceptional cases in which fully automated hardware warehouses are feasible). The use of retrieval agents that cooperate with the other agents, which are involved in the document management, gives in our opinion a solid base for an effective and efficient enterprise-wide information system.

3. A Cooperative Agent Architecture for Feedback

In order to find relevant information quickly it helps when documents contain meta information indicating their subject, intended audience and possibly other aspects. Unfortunately it is not possible to add meta information to external documents. So, we cannot assume that it is feasible to design and implement for every document an agent with knowledge about the document and its usage. What *is* feasible is that a user's retrieval agent is cooperating with a number of document warehouse agents that act as a kind of information brokers that know about the market place where documents are used (retrieved).

In the architecture that we propose here the learning retrieval agents add part of the document knowledge (meta-information) to their internal database (or user model). The internal database of an agent serves three purposes:

- The agent contains knowledge about the state of the process in which the user is involved. This knowledge is obtained through cooperating with the process agents.
- When a document is encountered again in a search the agent already knows whether the user finds it relevant or not under the given circumstances.
- When new documents have to be evaluated it helps to have a database with classifications of similar documents.

The best possible use of a retrieval agent's database is the help it can provide to other agents. When an agent encounters a document for the first time, other agents may already have classified that document. Although these other agents work for users with different interests, different jobs and tasks an agent can use the judgement of other agents in better evaluating a newly found document. We feel that summaries of this knowledge should be stored (learned) in special agents dedicated to the management of the document warehouse. These document warehouse agents can offer the common knowledge about the documents and their usage, and they can (on the basis of this knowledge) proactively control the contents of the document warehouse.

Cooperating agents are only feasible within a single organization, and most likely also only at a single site (or geographically near sites). This implies that agents do not have to make transformations between location and organizational information. (If one user has taught her agent that a document is relevant to her location, this applies to the other users' agents as well.)

Apart from reusing evaluations of documents from other retrieval agents, a retrieval agent may also ask other retrieval agents for documents about a certain topic and for a specific job and task. Only agents working for users with similar interests may offer help. Rather than simply reproducing these documents the agents also need to compare task information. When users have different jobs requiring different information, the agent needs the documents that were rejected by the other agent.

Altogether the architecture involves three types of cooperating agents:

- process agents: responsible for the flow of activities within the business processes, and thus responsible for the operational decisions involved in the execution of tasks

- document warehouse agents: responsible for the control of the document management process
- retrieval agents: responsible for the match between activities and their purpose on the one hand, and documents from the document warehouse on the other hand

4. Conclusions and Future Work

Information retrieval can be improved by separating knowledge about document content from the tasks a document is intended to support and the geographic location or organization it is aimed at.

By distinguishing three types of agents, process agents, document warehouse agents and retrieval agents, an organization can set up a retreival process in which the necessary knowledge is adequately distributed. When these agents cooperate, just like people cooperate in the traditional factory inventory processes, retrieval can be much better supported in flexible medium or large sized business environments.

Retrieval agents that assist users in finding information can help each other both by providing their evaluation of specific documents and by proposing documents based on the knowledge of each other's user model.

Evaluation of documents can be further enhanced by including a document's environment into that evaluation. Other documents pointing to a document, as well as pointers from the document under evaluation may provide valuable information about the purpose of a document. Also, the navigation path taken by a user to reach the document may provide cues about the type of information the user is searching for. These additional cues are not yet incorporated in our cooperating agents architecture, but will be in the near future.

5. References

[G97a] G. Golovchinsky.
Roll-Your-Own Hypertext. In *Proceedings of the Flexible Hypertext Workshop*, Macquarie Computing Reports C/TR97-06, pp. 49-53, 1997.

[G97b] G. Golovchinsky.
What the Query Told the Link: The integration of hypertext and information retrieval, In *Proceedings of the ACM Conference on Hypertext*, pp. 67-74, 1997.

[BL97] P. De Bra and W. Lemmens.
FishNet: Finding and Maintaining Information on the Net. (To appear) In *Proceedings of the AACE WebNet Conference*, Toronto, 1997.

[HD97] G.J. Houben and F. Dignum.
Information for organized work. In F. Baader, M. Jeusfeld, W. Nutt (eds), *Proceedings 4th Int. workshop Knowledge Representation meets databases*, Athens, 1997.

[MWFF]
Raúl Medina-Mora, Terry Winograd, Rodrigo Flores, Fernando Flores, The Action Workflow Approach to Workflow Management Technology, In *Computer-Supported Cooperative Work 92 Proceedings*, 1992, pg. 281-288.

MODELS, MEDIA AND MOTION: USING THE WEB TO SUPPORT MULTIMEDIA DOCUMENTS

DICK C.A. BULTERMAN

CWI: Centrum voor Wiskunde en Informatica
Kruislaan 413, 1098 SJ Amsterdam, Netherlands

The World-Wide Web has been used extensively to present hypertext documents that have a limited mixture of text and simple graphics which are distributed via the public Internet. The performance characteristics of the Internet have made the delivery of complex multimedia documents (that is, documents that include time-based components) difficult. An effort is currently underway by members of industry, research centers and user groups to define a standard document format that can be used in conjunction with time-based transport protocols over Inter- and Intranets to support rich multimedia presentations. This paper outlines the goals of the W3C's Synchronized Multimedia working group and presents an initial description of the first version of the proposed multimedia document model and format.

Introduction

The World-Wide Web is generally seen as the embodiment of the information infra-structure in today's information age. The Web's clever application of traditional technologies and its universal acceptance among a wide range of users has provided an information sharing backbone that is unique in history. The success of the Web is based largely on the use of a simple document format [8] and a straightforward (sub)document transfer protocol [9] . Using nothing more than a text editor and (if possible) an existing document as an example, even the most novice users can create complex hypertext documents which, using a fetch-and-store transport protocol, can be accessed from a variety of client computers across the Web.

The simplicity of the Web's structure is both a blessing and a problem. It has been a blessing because it has allowed a wide range of users to participate in the information infra-structure. Unfortunately, this simplicity has also limited the types of information that can be placed in Web documents. In a age where even low-end *PC* s have some support for audio and often video media, the Web has offered little or no support for fetching and displaying such time-based media items through standard document interfaces. *HTML* documents cannot express the synchronization primitives required to provide a coordinated presentation, and the *HTTP* protocol cannot provide the guaranteed delivery of time-based media objects required for continuous media data.

The development of Java extensions to *HTML* , known as Dynamic *HTML* [5] , provide one approach to introducing the necessary synchronization support into Web documents. This approach has the advantage that the author is given all of the control offered by a programming language in defining interactions within a document; this is similar to the use of the scripting language Lingo in *CD-ROM* authoring packages like Director [6] . Such an approach has the disadvantage that defining even simple synchronization relationships becomes a relatively difficult task for Web users who have little or no programming skills--the vast majority.

An alternative to using a programming language is the use of a declarative multimedia document format. In such a format, the control interactions required for multimedia applications are encoded in a text file as a structured set of object relations. The first system to propose such a format was *CMIF* [2] , [3] , [4] . Other more recent examples are *RTSL* [17] and *MADEUS* [10] .

In this paper, we describe a new declarative format for Web-based multimedia applications. This format is being developed by the Synchronized Multimedia working group of the *W3C* . While the development of the format is still in its formative stages, a review of the principles of this work can be useful to developers and researchers who are interested in the general direction of multimedia support for the *WWW* .

Section 2 presents background material that is useful in understanding the transfer of multimedia data in open networks. We begin with a short description of a typical Web-based application (the Web News), and then follow with a description of the infrastructure that can be used to actually transmit document components. This section closes with more background on the *W3C SYMM* working group. Section 3 provides an overview of the major aspects of the evolving format for describing multimedia applications. Here, we consider the encoding of temporal and spatial aspects of a presentation, as well as some rudimentary specification of alternate behavior based on characteristics of the presentation environment. We also consider the initial support planned for hypermedia aspects of documents. Section 4 closes with a discussion of open issues and related work.

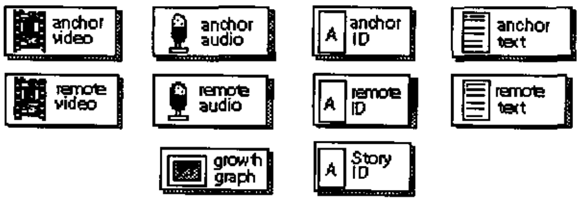
Web-Based Multimedia: Environment and Typical Applications

This section provides background information that will help define the types of applications and the support environment expected for first-generation Web-based multimedia documents. We begin with the description of a typical example, we then discuss the intended operational environment and close with a description of the *W3C SYMM* working group.

A Sample Document: The Web News

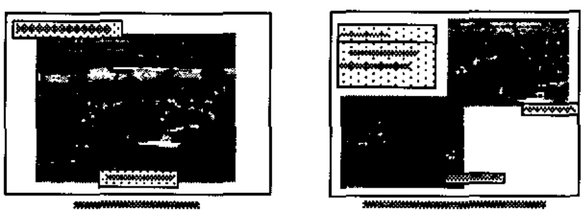
Multimedia applications have the general characteristic that they integrate a strong notion of time in a presentation. This is a sufficiently broad definition to encompass a wide range of applications, but it perhaps too broad to be of any great value for building a simple support mechanism for anything as anarchistic as the *WWW* . In order to focus our attention on the class of applications that this paper concerns itself with, we present an example of a generic Web application: that of a network newscast. 1

Several media objects can be defined that can make-up such a newscast. Let us assume that, for whatever reason, the objects shown in Figure 1 have been selected to make up the presentation. For purposes of this example, we do not care how each of these objects have been created, nor do we care where they are stored. What we do care about is that they have not been pre-packaged into a composite object that is fetched from a single source.

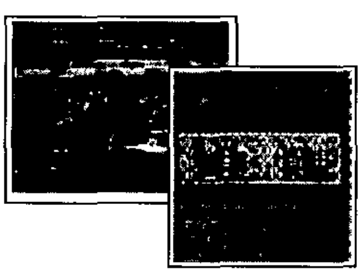


While storing the audio and video as separate objects will increase the synchronization burden on the playback environment, it increases the flexibility of the over-all presentation. One could conceivably substitute one audio track for another without having to rebuild the entire application. (This can be useful in multi-lingual environments.) Such substitution is not to be taken lightly, however, since often some form of content-based synchronization will may required among data objects.

Figure 2 shows two views of the newscast example, taken at different times in the presentation. On the left side, we see a portion of the introduction of a story on the growth of the World-Wide Web. In this portion, the anchor is describing how sales of authoring software are expected to rise sharply in the next six months. Figure 2 (b) shows a point later in the presentation, when the anchor is chatting with a remote correspondent in Los Angeles, who is describing how local Hollywood stars are already planning their own audio/video homepages on the Web.



The presentation described up to this point is dynamic in terms of content, but static in terms of structure--the entire presentation is played as if it were a single, composite object. A more interesting extension of this example is given in Figure 3 . Here we see the result of 'clicking' on the anchor: in addition to the presentation we first saw, an additional window has popped up which contains the anchor's home page. The ability to incorporate links to other pieces of content in the presentation transforms the static semantic structure in to a dynamic one, which is a powerful mechanism for creating complex presentations.



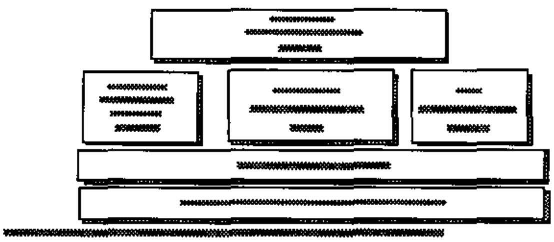
A total description of the implementation details of the Web newscast is beyond the scope of this paper. We will, however, refer to it as a running example in the sections below.

The Execution Environment

Presentations of the type discussed in the previous sessions can be transmitted over any type of network infrastructure, including the null infrastructure: *CD-ROM* . In a *CD-ROM* environment, it is possible to analyze standard system characteristics to determine the feasibility of presenting a document on a particular system. In a networked environment, this is often impossible. Especially in connectionless networks, authors cannot know ahead of time how many transport resources will be allocated to any one server-client data stream; this problem increases when each piece of data is saved as a separate object on separate servers.

In order to bring order into the chaos of sending time-restricted data over the Internet, the Internet Engineering Task Force *(IETF)* [7] has been developing a number of protocols that can serve to better manage the transfer of information between clients and servers. These include protocols for resource reservation, real-time transport, real-time control and real-time streaming of data. The deployment of some of these protocols is just starting to become a reality over *IP* -based public networks. While it is not clear if all of the protocols will be supported by all components of the network infrastructure--the universal acceptance of resource reservation seems doubtful--there is a concerted effort underway to augment the strict fetch-and-buffer approach used by *HTTP* .

Figure 4 shows the general relationship of these protocols to one-another. *IP* and *UDP* are standard protocols from the Internet suite. They are often implemented within the operating system kernel or as part of particular network device controllers. The other protocols are currently implemented as part of the application itself, rather than as part of the low-level operating system support. This is probably a transitional situation.



RSVP [12] , [13] is a resource reservation protocol that was designed for uni-directional multicast applications, but which also could be used for simplex unicast applications. As part of application initialization, a request can be made of intermediate network components to reserve resources at a particular quality of service (QoS) level for the lifetime of the application. From a Web perspective, the most interesting aspect of *RSVP* is that it is the first serious attempt at defining a network-wide reservations scheme. While it seems unlikely that *RSVP* will play a major short-term role in public networks, it may be a useful tool within Intranets for guaranteed end-to-end bandwidth allocation.

RTP [15] , [16] and *RTCP* [14] are the Real-time Transport Protocol and the *RTP* Control Protocol, respectively. Together they form a pair of transmission protocols that can serve as the basis for supporting time-based data delivery. While it may be natural to expect that a real-time protocol provides services to guarantee real-time (or on-time) packet delivery, this is not the case with *RTP/RTCP* . Instead, these protocols provide a framework and a set of building blocks with which a given application can create its own servicing algorithms to support on-time delivery of data packets. In this way, the particulars of the application and the data types being transferred can be used to determine the best support strategy, rather than relying on *RTP/RTCP* providing a 'one size fits all' type of service.

In practice, both *RTP* and *RTCP* make use of a local transport protocol to actually ship data between the source and destination(s) of a transfer. *RTP/RTCP support unicast and multicast transfers if the underlying transport mechanism does as well. RTP* is a packaging protocol that allows application data to be time-stamped and wrapped inside transport-level packets, such as those provided by UDP. From time-to-time, *RTCP* packets are sent between the source and destination(s) with transfer statistics. These can be used by the sender/receiver to adjust the manner in which data is buffered at either end of the transfer, or it can be used to dynamically select appropriate data encodings--but only if this is supported by the application itself.

RTSP [18] is a relatively recent streaming protocol that can be integrated with the protocols discussed above. (A streaming protocol is one that does not wait for an entire object to be delivered before rendering can begin.) As with *RTP* , *RTSP* does not a complete streaming solution; rather, it provides a framework in which applications-level streaming support can be implemented. To date, *RTSP* has been used commercially by Progressive Networks as part of their RealAudio/RealVideo suite [11] .

Having a set of transport and control protocols is an essential basis for supporting multimedia applications, even if these "protocols" provide only skeleton services. For the work described in this paper, these skeletons provide a common starting point.

Goals of the W3C Synchronized Multimedia Working Group

Until the beginning of 1997, there was no coordinated effort for non-proprietary multi- and hypermedia support through the World-Wide Web. In February of that year, the *WWW* Consortium (known as *W3C* [21]) initiated a working group on synchronized multimedia [22] . The *SYMM* working group was formed to develop a specification which carries the working title of *SMIL* , or the Synchronized Multimedia Integration Language [23] .

The development of *SMIL* grew out of a realization that a large class of hypermedia applications could (and probably should) be represented in a declarative format rather than as a computer program or high-level script. A declarative specification is often easier to edit and maintain than a program-based specification, and it can potentially provide a greater degree of accessibility to the network infrastructure by reducing the amount of programming required for creating any particular presentation. The success of *HTML* for hypertext documents has demonstrated the willingness and ability of Web users to create documents using a simple, structured format. In many ways, *SMIL* builds on this concept.

The *W3C SYMM* group has restricted its attention to the development of a common format, without specifying any particular playback or authoring environment. At present, approximately ten organizations have indicated an interest in developing prototype play-out environments. There also has been some support for developing or adapting authoring environments to generate *SMIL* encodings.

SMIL : Structured Document Format Specification

This section presents a summary of the proposed document format that is being developed by the *W3C SYMM* working group. We begin with a statement of the general characteristics of the format, and then consider the format's major components:

- temporal specifications: mechanisms to encode the temporal structure of the application and the refinement of the relative start and end times of events;
- spatial specifications: the primitives provided to support simple document layout;
- alternative behavior specification: the primitives to express the various optional encodings within a document based on systems or user requirements; and
- hypermedia support: mechanisms for linking parts of a presentation.

Each sub-section starts with a statement of general principles and then a description of the model components.

The goal of *SMIL* is to provide a declarative, text-based encoding of the behavior of hypermedia applications. Once encoded, the document should be able to be played on a wide-range of *SMIL* browsers. Such browsers may be stand-alone presentation systems that are tailored to a particular user community or they could be integrated into standard browsers.

Media Objects

Perhaps the most fundamental aspect of *SMIL* is that it is an integrating format. Unlike *HTML* , an *SMIL* document contains only structure and media object description information--it does not contain any data associated with the objects themselves. The display software for *SMIL* documents (either stand-alone players or adapted browsers) must be able to render the individual data components based on the description in the *SMIL* file. The use of an integrating format is essential for multimedia applications: unlike text, even the most trivial audio or video object can contain massive amounts of data. Storing these items within the document would make its size unmanageable.

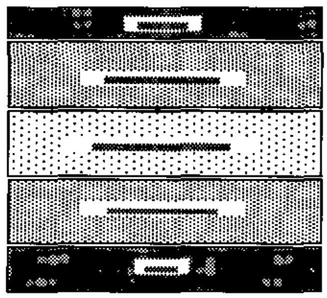
References to individual media objects are made via media object instance specifiers. These are of the general form:

<{type} SRC="{protocol}:{location}/{name}" {attributes}>

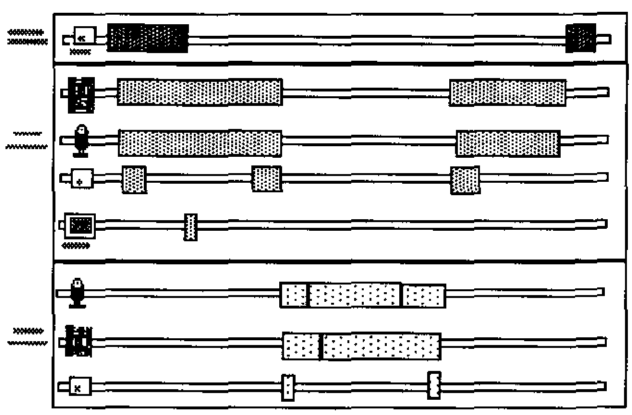
where *type* indicates the type of data (such as text, image, video, etc.), the *SRC* field provides the familiar object reference, and the *attributes* fields provide details that are discussed below.

Temporal Specifications

If we were to define the Web Growth story outlined in Section 2.1 in terms of its over-all structure, we would wind up with a representation similar to that in Figure 5 . Here we see that the story starts with a opening sequence (perhaps containing a logo and a title), and ends with a similar closing segment. In between is the "meat" of the story. It contains an introduction by the local anchor, followed by a report by the remote correspondent, and then concluded by a wrap-up by the local anchor. This 'table of contents' view defines the basic structure of the story. It may be reusable, in that many stories may be similarly structured.

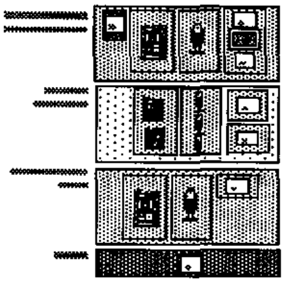


This structure view, without any references to particular media objects, does not provide sufficient detail to describe an example, but it can be used to define a general sequence of story elements. A more common view of an application is shown in Figure 6 , where we see a timeline of the Web Growth story. Rather than illustrating structure, this view shows each of the components and their relative start and end times. (Note that, as part of the initial anchor setup, a reference is made to a graph that is active during only a part of the presentation.) While timelines provide an effective graphical representation of an application--so long as that application is not too complex--it is not a useful basis for encoding an application in a portable manner.



The timeline view does provide an insight into the actual temporal relationships among the elements in the presentation. Where Figure 5 assumed that the opening segment occurred entirely before the first anchor segment, we see that there is actually some overlap of the opening text and the anchor audio and video. The remaining structural partitioning remains correct, although it is clear that within each element a number of media events occur in parallel or sequentially.

If we were to combine the structure and timeline views, we might end up with the representation shown in Figure 7 . This representation shows that the story is made up of a number of events that occur sequentially or in parallel. Within the parallel events, some start at the exactly the same time, while others start at an offset relative to each other.



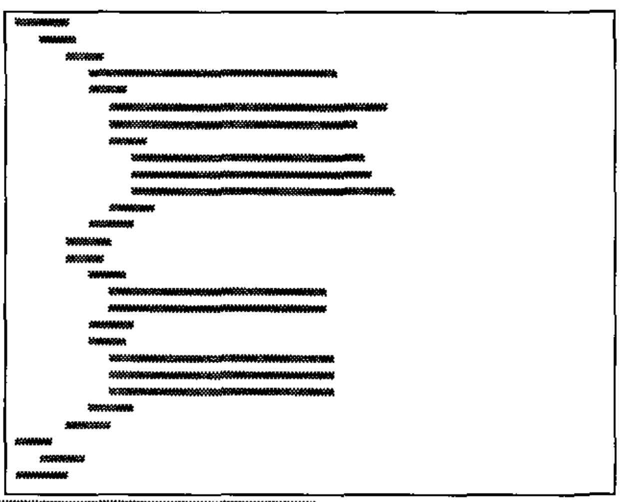
As with the timeline, a structure diagram is an unhandy way to encode a portable document. Instead, the *SMIL* format uses the following two structuring elements, taken from *CMIF* :

<seq> ... </seq> : A collection of objects that occur in sequence.

<par> ... </par> : A collection of objects that occur in parallel.

Elements defined within a *<seq>* group have the semantics that a successor element is guaranteed to start after the completion of a predecessor element. Elements within a *<par>* group have the semantics that, by default, they all start at the same time. Once started, all elements are active for the time determined by their encoding or for an explicitly defined duration. Elements within a *<par>* group can also be defined to end at the same time, either based on the length of the longest component or on the end time of an explicit master element. Note that if objects within a *<par>* group are of unequal length, they will either start or end at different times, depending on the attributes used to define the group.

The structural elements can be nested to describe applications of arbitrary complexity. A partial encoding of our example story is shown in Figure 8 . (Note that the syntax of the object references is generalized to improve readability.) The encoding shows that the Web Growth story contains a sequence of parallel groups, some of which contain nested *<seq>* and *<par>* elements. In the first *<par>* , we see that the text object story_heading is presented in parallel with the initial anchor setup. This setup consists of an audio and video stream that is played in parallel, along with a sequence of media objects: a text label, followed by an image and then a text label.



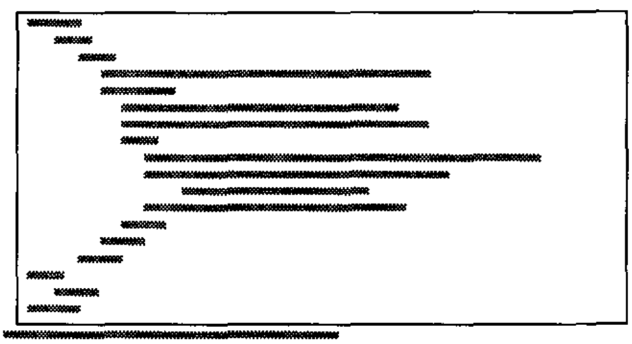
If we compare Figure 8 with Figure 7 , we see that the *SMIL* encoding as it is presented gives only the coarse structure of the application's temporal relationships. For example, the actual overlap between the story_heading and the anchor audio/video tracks is minimal. The timeline also shows that the display of the label anchor_name is to happen shortly after the anchor appears in view. Also, the image web_growth in the nested <*seq*> group needs to appear on the screen when the anchor refers to it in the story. To handle these types of situations, *SMIL* provide three types of timing control relationships:

- explicit durations: a *DUR=*" time " attribute can be used to state the presentation time of the object; 2
- absolute offsets: the start time of an object can be given as an absolute offset from the start time of the enclosing structural element by using a *BEGIN=" time "* attribute;
- *relative offsets* : the start time of an object can be given in terms of the start time of another sibling object using a *BEGIN=" object_id + time "* attribute.

(Unless otherwise specified, all objects are displayed for their implicit durations--defined by the object encoding or the length of the enclosing <*par*> group.)

The specification of a relative start time is a restricted version of *CMIF* 's sync_arcs [4] to define fine-grain timing within a document. At present, only explicit time offsets into objects are supported, but a natural extension is to allow content markers, which provide content-based tags into a media object.

If we use the attributes defined above, the initial part of the application's encoding can be re-written as illustrated in Figure 9 .



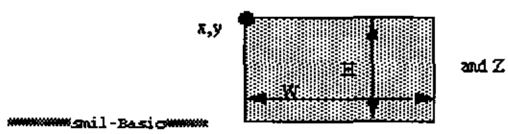
Layout Specifications

Since *HTML* is based on a text flow model, one of its advantages is that an author is able to define a presentation without worrying about the exact positioning of individual objects in that presentation. The actual "look and feel" of the document was determined at run-time depending on the screen space allocated to the browser and on user preferences. This de-coupling of content and presentation led many in the *SYMM* group to consider presentation layout to be beyond the scope of the format. While some relationship does exist between the timing hierarchy and the ultimate presentation layout, it became clear that, for multimedia applications, extra facilities were required to determine the relative positioning of media objects.

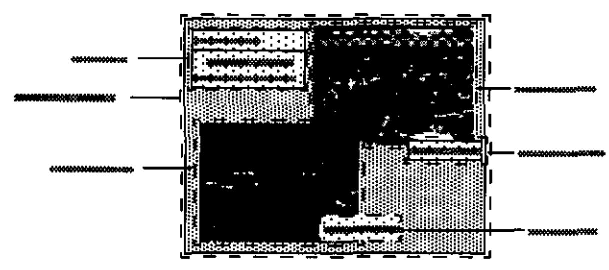
A generalized solution to presentation layout was still considered outside the scope of the format, but it has been agreed that, to support some form of inter-operability among the potential variety of *SMIL* players, some rudimentary form of layout control is required. To this end, it was decided that four types of layout schemes would be supported by the format:

- the null (default) layout: a layout scheme that is appropriate for very simple documents (those containing one video and one audio, each of which consume all of the available resources);
- a bare-bones basic layout: basic positioning is supported, without trying to solve the general layout problem;
- a hook for external layouts, with *SMIL* as the master: an *SMIL* document forms the basis of a presentation, but which uses facilities available in a particular player;
- a hook for external layouts, with *SMIL* as the slave: an external player starts an *SMIL* document as part of a larger presentation.

It was further agreed that all *SMIL* renderers should support the same semantics for default and basic layout--thus providing a means for maintaining compatibility--and that each renderer would also have the option to support its own master or slave layout semantics. The creator of an *SMIL* document would be free to choose the layout scheme which met the application's needs.



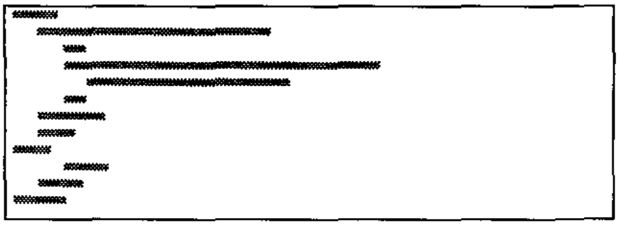
The general capabilities of the *smil-basic* layout specification is given in Figure 10 . Each visible or audio object is rendered to a layout channel. (These channels have their roots in similar constructs used in *CMIF* and *RTSL* .) A channel is an abstract entity that can describe screen space, audio channels, or any other rendering resource. Associated with a screen channel are x, y coordinates that define the upper-left anchor point of the channel, plus a *HEIGHT* and *WIDTH* , measured in pixels or percentages. Channels also have an integer *Z* depth associated with them, with greater values indicating objects that are closer to the viewer. As such, the channel is a reusable, indirect reference to a set of presentation coordinates. Figure 11 shows the channel map associated with a part of the Web newscast. This portion has six visible channels (plus associated audio channels). The presentation of all channels is managed by the renderer.



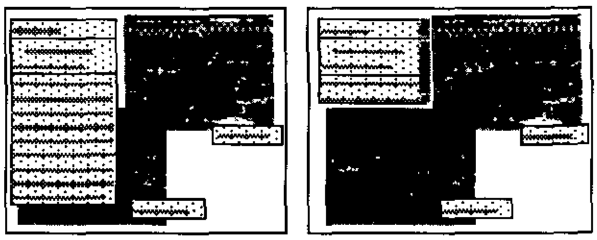
Each media object instance contains a channel reference:

<video href="rtsp:anchor.mpg" channel="anchor">

which refers to a similar tag in a layout specification. The layout specification itself has the form shown in Figure 12 . In order to accommodate systems that have richer (or poorer) layout semantics than *smil-basic* , an alternative layout definition scheme is provided. This is discussed in Section 3.4.



If a particular media object has a "natural" height and/or width, these values can be left unspecified. Care must be taken to make sure that objects of undefined size are bound to a particular screen area; if not a situation such as that illustrated in Figure 13 could occur.



Alternate Behavior Specifications

If the presentation infrastructure for an *SMIL* file were known before-hand, one file could be tailored to that environment easily. Unfortunately, this is nearly never the case in the Web. Clients will access a document from a wide range of locations, and will encounter a wide range of transmission and server delays.

Providing truly adaptive documents, that match their performance and appearance characteristics to the resources available is a fascinating research topic. (It is fascinating in large part because it is unsolved in the general case.) In the context of *SMIL* it was felt that some form of support for adaptive behavior was required, even for simple first-generation documents.

The solution adopted by the SYMM group was to provide a means for defining alternate behavior within a document, but to rely totally on the presentation environment to resolve which of the alternatives would be selected at run-time. This selection could take place based on profiles, user preference, or environmental characteristics. Clearly, specifying this type of behavior within the confines of an adaptive format is a challenging proposition.

As an example of specifying alternate behaviors, consider the following:

<switch>

<audio profile=bad_uk src=low-res.aiff />

<audio profile=good_uk src=hi-res.aiff />

<audio profile=bad_nl src=low-res.aiff />

<audio profile=good_nl src=hi-res.aiff />

</switch>

In this fragment, we assume the presence of four alternate audio files. Two of these files contain English-language audio, and two contain Dutch-language audio. A high resolution and low resolution version is available for each language. At run-time, the player could evaluate the alternatives based on its own algorithms and select the alternative that is most appropriate. Since all choices are semantically equal, any particular item can always be considered to be a "correct" choice (if perhaps not optimal); as a result, a player could always choose to select the first alternative if it wished.

An example of the application of alternative is the use of multiple layout specifications in a document. Suppose that an author wanted to guarantee some measure of compatibility across all players, but that s/he also wanted to exploit the fancy features offered by a particular environment. In this case, the following specification could be used:

<switch>

<layout type="text/smil-basic">

. . .

</layout>

<layout type="text/cmif">

. . .

</layout>

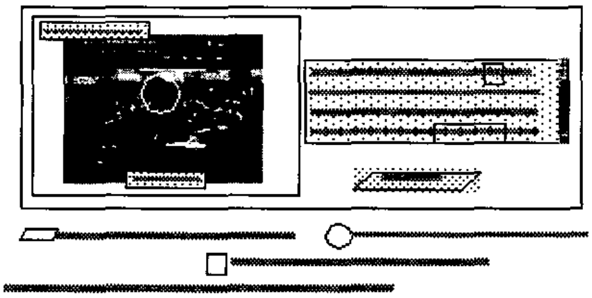
</switch>

If the application was played on a non-CMIF player, then it could choose to use the *smil-basic* layout features. If it was played on a CMIF player, then CMIF's multi-window functionality could be exploited.

Hypermedia and *SMIL*

As the current Web has demonstrated so effectively, a networked information infrastructure is based in large part on its ability to reference related pieces of information from within a presentation. With *HTML* , this is a straight-forward process: each document has a single focus (the browser window or frame) and anchors and links can be easily placed within the document text. In an *SMIL* presentation, the situation is much more difficult. First, the location of a given anchor may move over time--and even if it does not move, it still may be visible for only part of the object's duration. Second, since SMIL is an integrating format, conflicts may arise on ownership of anchors and the semantics of following any given link.

As an example of the problems that can be encountered with links, consider the situation outlined in Figure 14 . Here we see three visible channels, one containing an embedded *SMIL* presentation, and two containing *HTML* text. The top *HTML* text is a conventional page (with internal links) while the bottom *HTML* text (labelled next story) contains an *SMIL* link. Since *SMIL* integrates many presentation data types, it is important to know if following a link from any given anchor will result in intra-object or inter-object navigation. We can identify three situations in which a link can be defined: as a link defined by the containing *SMIL* document, as a link defined by an embedded *SMIL* document--that is, an *SMIL* sub-document, or a link defined within a non- *SMIL* component.



If the link is defined by a containing *SMIL* document, such as the next story link, activation of the link affects the presentation of the whole *SMIL* document. This effect depends on the value of a *SHOW* attribute, which may have values:

- *REPLACE:* the presentation of the destination resource replaces the complete, current presentation (this is the default);
- *NEW:* the presentation of the destination resource starts in a new context (perhaps a new window) not affecting the source presentation; or
- *PAUSE* : the link is followed and a new context is created, but the source context is not replaced but is suspended.

If the link is defined by an embedded *SMIL* document, such as by following the link placed over the head of the news anchor, activation of the link affects only the embedded *SMIL* document. The effect depends on the value of the *SHOW* attribute as described above.

If the link is defined by an non- *SMIL* document which is embedded in an *SMIL* document, such as following one of the references in the text in the upper right of the figure, link traversal can only affect the presentation of the embedded component and not the presentation of the containing *SMIL* document. This restriction may be released in future versions of *SMIL* .

General support for hypermedia is a complex task. Interested readers are invited to study the approaches defined for the Amsterdam Hypermedia Model [3] , which serves as the basis for the current *SMIL* proposal.

Current Status and Future Directions

As of the writing of this article (late September, 1997), may of the details of the encoding of *SMIL* were being finalized. For the latest version of the specification, interested readers should consult *W3C* 's web pages.

As part of the *W3C* approach to defining Web standards, the format will be submitted to interested parties, who can choose to develop prototype implementations. (At present some eight organizations are involved in developing prototype environments.) These implementations will be evaluated against a set of standard application examples to determine the viability of the encoding format. The format--plus changes--will then be submitted to the full membership.

While the specific choices made in developing *SMIL* are interesting in their own right, perhaps the most interesting aspect of the format is that is can provide a common base for future research on various aspects of networked multimedia systems. Where in the past true multimedia applications were could never inter-operate with research platforms, *SMIL* provides a foundation that will allow comparative examples to be developed that can support experimental research. If this can be accomplished, then the work put into the development of the specification will be worth the considerable effort it has required.

The work of the *W3C* 's *SYMM* working group was coordinated by Philipp Hoschka of *W3C/INRIA* . A complete list of contributors is available at [22] .

CWI 's activity in this project is funded in part through the *ESPRIT-IV* project C *HAMELEON* [1] of the European Union. Additional sources of funding have been the *ACTS SEMPER* [19] and the Telematics *STEM* [20] projects.

In the spirit of Web-based information exchange, references to publications have been given to reports available via Web servers as much as possible.

1. CHAMELEON: An Authoring Environment for Adaptive Multimedia Presentations. ESPRIT-IV Project 20597. See *http://www.cwi.nl/Chameleon/.* See also *http://www.cwi.nl/~dcab/.*
2. D.C.A. Bulterman, R. van Liere and G. van Rossum: A Structure for Transportable, Dynamic Multimedia Documents (with R. van Liere and G. van Rossum), Proceedings of 1991 Usenix Spring Conference on Multimedia Systems, Nashville, TN, 1991 pp. 137-155. See also *http://www.cwi.nl/~dcab/.*
3. L. Hardman, D.C.A. Bulterman and G. van Rossum: The Amsterdam Hypermedia Model: Adding Time and Context to the Dexter Model-CACM 37(2), Feb. 1994, pp 50-62. See also *http://www.cwi.nl/~dcab/.*
4. D.C.A. Bulterman: Synchronization of Multi-Sourced Multimedia Data for Heterogeneous Target Systems, in Network and Operating Systems Support for Digital Audio and Video (P. Venkat Rangan, Ed.), LNCS-712, Springer-Verlag, 1993, pp. 119-129. See also *http://www.cwi.nl/~dcab/.*
5. W3C: HTML 4.0: W3C's Next Version of HTML. *http://www.w3.org/MarkUp/Cougar/*
6. Macromedia, Inc.: Director 6.0. *http://www.macromedia.com/software/director/*
7. IThe Internet Engineering Task Force: Home Page. *http://www.ietf.cnri.reston.va.us/*
8. W3C: HTML Version 3.2. See *http://www.w3.org/MarkUp/.*
9. W3C: Jigsaw Overview. See *http://www.w3.org/Jigsaw/*
10. C. Roisin, M. Jourdan, N. Layaïda and L. Sabry-Ismail: Authoring Environment for Interactive Multimedia Documents, Proc. MMM'97, Singapore. (Elsewhere in these proceedings.) See also *http://opera.inrialpes.fr/OPERA/multimedia-eng.htm* l
11. PRealNetworks (formerly Progressive Networks). Home page.

See *http://www.real.com/*
12. Zhang, L., Braden, R., Estrin, D., Herzog, S., and S. Jamin, Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification, *ftp://ftp.ietf.org/internet-drafts/draft-ietf-rsvp-spec-16.ps* .
13. Zhang, L., Deering, S., Estrin, D., Shenker, S., and D. Zappala, RSVP: A NewResource ReSerVation Protocol, *ftp://parcftp.xerox.com/pub/net-research/rsvp.ps.Z* , IEEE Network, Vol 7, no 5, pp 8-18, September 1993.
14. D. Bettati, et al: Connection Establishment for Multi-Party Real-Time Communication, Proc. NOSSDAV '95, Durham NH 1995. See *http://www.cs.berkeley.edu/~amit/research/Postscript/estab-nossdav.ps*
15. H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, rfc1889: RTP: A Transport Protocol for Real-Time Applications, 01/25/1996. *http://www.cis.ohio-state.edu/htbin/rfc/rfc1889.html.*
16. H. Schulzrinne, RTP, *http://www.cs.columbia.edu/~hgs/rtp/.*
17. RTSL: Real-Time Streaming Language. See *http://www.real.com/server/intranet/index.html*
18. H. Schulzrinne: A real-time stream control protocol (RTSP), *http://www.cs.columbia.edu/~hgs/rtsp/draft/draft-ietf-mmusic-stream-00.txt*(11/26/96).
19. SEMPER: Secure Electronic Marketplace for Europe, EU ACTS project 0042, 1995-1998. See *http://www.semper.org/* .
20. STEM: Sustainable Telematics for the Environment, EU Telematics Project EN-1014, 1995-1996.
21. W3C: The World-Wide Web Consortium. *http://www.w3.org/* .
22. W3C Synchronized Multimedia Working Group. *http://www.w3.org/AudioVideo/Group/* .
23. W3C SMIL Draft Specification. *http://www.w3.org/AudioVideo/Group/WD-symm-format.html* .

---

1. In 1991, as part of the first public CMIF paper, a network newscast was also used as a prototypical example. Sometimes the electronic world moves as slowly as the real one!

2. The syntax used to express time is still under discussion within the group.

Cartografische Agenten voor Visualisatie bij Geografische Informatiesystemen.

Robert E. Kuunders

GIS-Technologie TUDelft

R.E.Kuunders@geo.tudelft.nl

Inleiding

In deze samenvatting wordt aangegeven hoe gebruik wordt gemaakt van Geografische Informatiesystemen voor de analyse en weergave van ruimtelijke informatie. De inzet van technieken die goed toepasbaar zijn bij gedistribueerde gegevens zoals bij internet blijken ook toepasbaar bij de omgang met ruimtelijke informatie. Omdat een browser of een hierop gebaseerd platform uitstekend is uitgerust om gedistribueerde gegevens én applicaties te combineren lijkt het een geschikt medium om gedistribueerde ruimtelijke gegevens te combineren en te visualiseren met behulp van gedistribueerde applicaties.

Geografische Informatie en Informatiesystemen

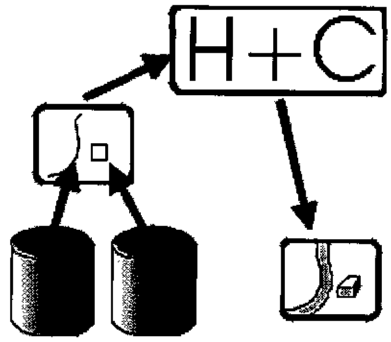
Een groot aantal disciplines in de wetenschap houdt zich bezig met processen en objecten die zich in, op en rond onze aarde afspelen. Deze processen en objecten hebben als altijd aanwezig kenmerk de positie in de ruimte, gemeten naar een gemeenschappelijk referentiekader. Daar de ruimtelijke positie bij deze processen en objecten een centrale rol speelt worden aan de informatiesystemen waar de gevens in worden opgeslagen en verwerkt ook bijzondere voorwaarden gesteld. Een Geografische Informatiesysteem (GIS) heeft dan ook als kenmerk dat aan alle objecten coördinaten zijn toegekend en dat met behulp van de overige attributen van de objecten ruimtelijke relaties kunnen worden opgeslagen, geanalyseerd en weergegeven. Door gebruik te maken van een bekend referentiesysteem (of beter coördinaatstelsel) kunnen andere geografische gegevens worden toegevoegd en betrokken in de ruimtelijke analyses.

Door de ruimtelijke component is het zoeken en onttrekken van gegevens aan een GIS een gespecialiseerde taak. Anders dan bij bijvoorbeeld bibliotheeksystemen kan niet worden volstaan met een trefwoordenregister en een alfabetische index. Typische vragen bestaan niet alleen uit de vraag 'waar' maar tevens uit de vraag 'in relatie tot'; "Ik zoek een wandelend bereikbare telefooncel dichtbij mijn huis."

De weergave van de ruimtelijke gegevens, aanwezig in een GIS, geschiedt doorgaans door middel van kaarten. De communicatieve waarde is sterk afhankelijk van de mogelijkheden die het GIS biedt en de cartografische kennis van de operateur. Ervaring heeft geleerd dat vrijwel alle bestaande GIS niet goed in staat zijn goede cartografische producten te leveren zonder intensieve nabewerking.

De inzet van Geografische Informatiesystemen ten behoeve van Cartografische Visualisatie

De toename van het gebruik van GIS als hulpmiddel bij planning, ontwerp en verantwoording van ruimtelijke ingrepen zorgt voor een toegenomen behoefte om de resultaten toegankelijker te maken voor een breder publiek. De communicatieve waarde van de resultaten moet worden verhoogd waarbij tevens een streven bestaat het grootste deel van het proces zo geautomatiseerd mogelijk te laten verlopen. De weergave van de uitkomsten van ruimtelijke analyses kan op verschillende manieren, van traditionele papieren media via beelscherm en luidspreker tot volledige Virtual Reality omgevingen. In alle gevallen worden er specifieke eisen gesteld aan de weergave, afhankelijk van het medium.

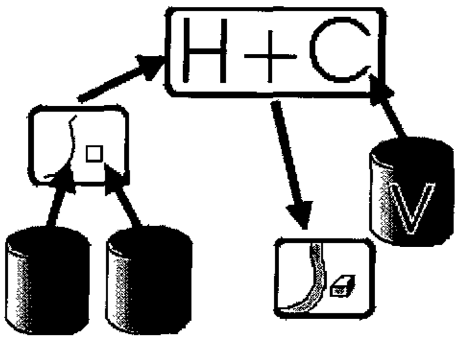


Traditionele totstandkoming van een cartografische afbeelding bij een GIS. Door ingrepen van de mens (H) en de hulp van een computer (C) komt de uiteindelijke weergave tot stand.

De cartografie houdt zich reeds decennia bezig met de afbeelding van geografische gevens. De ontwikkeling van computertechnologie heeft ook in deze tak van wetenschap belangrijke veranderingen met zich meegebracht. Met name de mogelijkheden van weergave (en communicatie) van ruimtelijke structuren en processen zijn verruimd van het platte papieren vlak tot multimedia- en hypermedia-technieken inclusief Virtual Reality. In de cartografie zijn een aantal stromingen onstaan die op verschillende wijzen gebruik maken van moderne technieken en de rol van de cartografische producten een andere invulling geven. Er treed een verschuiving op van een vooraf bepaalde communicatieve waarde naar een meer vrijere op wetenschappelijke visualisatie lijkende vorm van cartografie. Dit heeft tot gevolg dat de uiteindelijke cartografische weergave minder gebonden hoeft te zijn aan cartografische regels, maar dat tegelijkertijd de eindgebruiker een grotere eigen verantwoordelijkheid heeft om de (meest) juiste boodschap aan de afbeelding te onttrekken. Een bijkomend voordeel is dat juist die grotere vrijheid bij de samenstelling van de cartografische afbeelding de ruimte geeft om verdere automatisering door te voeren. Zeker wanneer resultaten worden getoond op een tijdelijk medium zoals een computerscherm en tevens een strategie van alternatieven kan worden gevolgd is de inzet van agent technologie een reëele optie in de vorming van een cartografisch product.

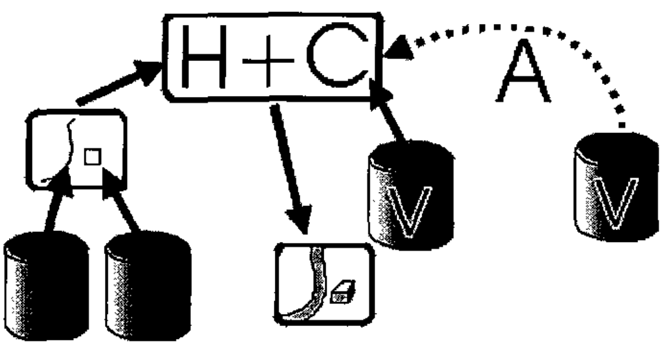
Wanneer men wil beschikken over alle geografische gegevens van een bepaald gebied, bijvoorbeeld Nederland, dan zal men daar een grote hoeveelheidheid opslagruimte voor beschikbaar moeten stellen. Het streven is dan ook niet gericht in de creatie van één grote gegevensbank, maar in de uitwisselbaarheid van de gegevens van meerdere gegevensbanken. Zeker gezien de toename van de hoeveelheid geografische gegevens, de inwinning verloopt ook steeds meer geautomatiseerd, is het een noodzaak dat verschillende bestanden aan elkaar gekoppeld kunnen worden en dat (deel)resultaten van analyses kunnen worden uitgewisseld.

Voor het uitvoeren van een ruimtelijke analyse worden eerst de betreffende gegevens verzameld en vervolgens gecombineerd in één gegevensbank om daarna de analyse uit te voeren. Er zijn suggesties om dit proces in de toekomst anders te laten verlopen. In plaats van de gegevens te kopieren en weer opnieuw op te slaan kan men ook het gebruik van de gegevens via een netwerk toestaan, zonder dat deze direct opnieuw moeten worden opgeslagen. Bij de analyse worden dan alleen die gegevens betrokken die een rol spelen en de eigenaar van de gegevens hoeft niet de hele gegevensbank af te staan voor het gebruik van slechts enkele gegevens. Niet de gegevens worden dus uitgewisseld maar een deel van de analyse. De resultaten van deze deelanalyses worden verzameld en gecombineerd bij de eindgebruiker.



Naast de vaste computermogelijkheden kan tevens een database met visualisatiegegevens geraadpleegd worden. Hierdoor ontstaat al een grotere vrijheid en variëteit van weergave.

De weergave van de gegevens en de verkregen resultaten zou op eenzelfde wijze kunnen verlopen. Naast of in plaats van geografische gegevens kan men ook een wijze van weergave opnemen in een gegevensbank. Deze informatie zou tevens in de begeleidende informatie bij ruimtelijke gegevens (de zogenaamde Meta-Informatie) kunnen worden opgenomen. De eindgebruiker zou vervolgens via een keuze in de legenda (het verklarende deel van een kaart) kunnen aangeven welke vorm van weergave de voorkeur heeft of een reeds bepaalde voorkeur voorgeschoteld kunnen krijgen. Er zouden dan naast de bestaande clearinghouses voor geografische informatie ook clearinghouses met weergave informatie kunnen worden ingericht. Op deze plekken zouden agenten gewezen kunnen worden op bronnen die weergave informatiebevatten en elkaar kunnen leren waar bepaalde informatie voorkomt en hoe verschillende weergavevormen worden gewaardeerd door eindgebruikers. Deze kennels voor agenten vormen een kennisstructuur bovenop de datastructuur van de geografische gegevens.



Wanneer visualisatie met behulp van agenten plaatsvindt is de vrijheid tot experimenteren groter terwijl toch vertrouwd kan worden op ervaringen en adviezen van deskundigen die de gegevens aan agenten ter beschikking stellen.

Discussie

Het ter beschikking stellen van specialistische kennis kan ook nadelig werken. De keuze neemt toe waardoor ook de mogelijkheid tot verwarrende presentatie van de gegevens toeneemt. Het is belangrijk te beseffen dat bij weergave van thematische informatie toepassing van kennissystemen vrijwel onmogelijk is, tenzij goed te voorspellen valt wat voor informatie wordt weergegeven. Bij een vrije keuze van exemplarische informatie voor weergave kan het best worden uitgegaan van standaard (of gestandaardiseerde) elementen die vaak in kaarten voorkomen. Hieronder vallen gegevens zoals topografie, weer/klimaat en landbedekking. De ervaringen hiermee opgedaan kunnen gebruikt worden om meer thematische onderwerpen zoals demografische gegevens op te nemen. Er zijn reeds diensten en bedrijven zoals RAVI (Raad voor Vastgoed Informatie) en de Topografische Dienst die overwegen voorbeelden van weergave van door hen geleverde gegevens ter beschikking te stellen.

Hypermedia Style Sheets on the World Wide Web

Jacco van Ossenbruggen (*), Lynda Hardman (**), Lloyd Rutledge (**)
(*) Vrije Universiteit, Fac. of Mathematics and Computer Sciences
De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands
(**) CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
email: jrvosse@cs.vu.nl

October 27, 1997

1 Introduction

In the first years after the introduction of the Web, document modeling development focussed on the *appearance* rather than on the content of HTML pages. A plethora of new tags were added to the original HTML specification and, to further enhance the layout capabilities, Java applet support and plug-ins became standard features of Web browsers.

More recent developments focus on the *structure* of Web pages and more importantly, on the separate definition of the structure (in the document itself) and its layout (in a style sheet). As a consequence, we can now separate layout from an HTML page by using the Cascading Style Sheets (CSS1 [5]) W3C Recommendation, which has been implemented by all major browser vendors.

HTML markup, however, remains oriented towards physical appearance and has limited facilities for a richer structural markup, which is needed to convert the displayed content of today's Web pages to machine-readable and interchangeable structured content [4]. The eXtensible Markup Language (XML [2]) provides a standard, open format to define such documents with a rich and machine readable structure. Because a browser has generally no prior knowledge about the tags in an arbitrary XML document, it depends completely on an associated style sheet which defines how to process the markup and content of the document. This requires a style sheet language which is much more powerful and expressive than CSS1. The eXtensible Style Language (XSL [1]) provides such a language for page-based documents and is based on the ISO standard DSSSL [3].

In this paper, we claim that for true hypermedia presentations we need a style language with a number of significant extensions to the current XSL proposal. We discuss the existing use of style sheets as applied to text and discuss their application to the case of hypermedia, and in particular how they need to be extended. A preliminary version of this paper was presented in [6].

2 Use of Style Sheets in Hypermedia

From an abstract point of view, style sheets specify the mapping of a model describing the logical structure of the document onto an output model. The output model may also be a logical structure model, or may describe a presentation structure. For page-based documents (e.g. traditional print or HTML) a clear abstraction of such a presentation structure has been developed. This is based on a sequence of pages consisting of columns, headers and footers, etc. [3]. This abstraction allows the specification of style sheets which map a variety of logical document models (including encyclopedia, novels, articles and letters) onto the same presentation structure — this technique has long been commonplace.

Style sheets for traditional printed documents associate individual document structures with presentation attributes of the medium such as font type, character style and font size, and also describe aspects inherent to the presentation structure, such as page size, margins and headers. Work has been carried out on style sheets for HTML documents on the World Wide Web [5]. HTML is an extension of text including anchors and links, and as such the style sheet options have been extended to include the visual styles of anchors.

In order to take style definitions further for application to hypermedia we need to include not only the anchor and linking aspects of hypertext, but also the structural and scheduling aspects of multimedia. The most important of these is the temporal specification of the document, given that spatial aspects have been partially addressed for the text case in terms of pages.

To illustrate possible uses of a hypermedia style sheet we present a simple scenario. A Web document describing a slide show as an ordered collection of slides may be presented in several ways. For example, each slide can be presented for a fixed time interval (e.g. 10 seconds). Here the style sheet need only associate a duration with each slide. An alternative manner of presentation is that the slide is left on display until the user selects a link to the following slide. Here the style sheet has to specify the addition of a hyperlink for each slide plus some means of activating the link, for example introducing a "next" text button. For a more advanced audience the author may wish that a number of slides are skipped, or the ordering changed. The style sheet would specify the structural changes and the output document could then be given presentation characteristics by either of the previous two style sheets.

3 Extensions to Hypermedia Style Sheet Models

To generate a presentation, a mapping from a hypermedia document onto a presentation needs to be specified. This mapping includes the specification of temporal and spatial constraints, and the specification of hyperlink behavior in documents containing multiple synchronized media streams. The currently available style sheet languages, and the models on which these languages are based, are currently unable to support this mapping. The following gives a brief characterization of these problems (which will be discussed in the full version of this paper in more detail).

Temporal Constraints In contrast to the two dimensions of text, hypermedia contains a third, temporal dimension. This dimension has to be reflected in the final presentation to ensure the precise scheduling and synchronization of the media items involved. Often, the logical document structure also reflects some of the important temporal constraints as well, i.e. hypermedia authors tend to organize their documents in a way that reflects the temporal structure. Examples include the logical grouping of elements which have to be played in sequence or in parallel. Style sheet languages do not support the specification of temporal constraints because these cannot be expressed in their output models.

Spatial constraints In contrast to spatial layout based on text-flow used in word-processors and hypertext browsers, hypermedia applications require different mechanisms to specify their spatial layout. In general, resizing the browser's window should not result in re-formatting the document to generate a new text-flow, but result in resizing of the media items playing. This may involve constraints to preserve aspect ratios or cropping of media items which do not support scaling. Because the output models of most style sheet languages are based on a flow model, the layout requirements of hypermedia documents cannot be specified by these languages.

Hyperlinking continuous media objects Continuous media objects introduce additional complexity to the specification of hyperlinks among these objects. Anchors may be active only for a certain amount of time, or be connected to different links during different time intervals. Video anchors probably move during the presentation.

Additionally, link traversal might lead users to a point somewhere in the middle of another document, as is the case in HTML. For hypermedia documents, this might involve fast-forwarding the presentation to start it at the right moment. Because their output models do not account for temporal constraints, this type of link behavior cannot be specified by current style languages.

Hyperlinking within multiple active media streams In hypertext, a window usually displays a single document (HTML frames are a common exception) where, on traversal, the destination of a link replaces the complete source document, or is displayed in a new window. In hypermedia documents, multiple streams of media items might be active simultaneously, and link traversal should not necessarily affect all of them. As a consequence, links have to define what their context is terms of their source (i.e. which of the currently active streams are affected) and in terms of their destination (i.e. which of the streams of the destination will be activated). Synchronization constraints between objects belonging to the streams involved might further complicate the link processing. The facilities for defining link behavior are very limited in most style sheet languages, and few models support the concept of link context.

Style sheets in adaptive environments In adaptive environments, the style sheet conversion might also be based on information about how to adapt to user characteristics (e.g. level of expertise) or changing system resources (e.g. network bandwidth). The hypermedia style sheet could be used to indicate how to deal with limited resources (by specifying alternatives or QoS negotiation protocols) on different platforms, thus making the document source independent of platforms specific details.

4 Conclusions and Future Work

Style sheets are already a standard feature of text processing systems. For hypermedia systems to benefit from the use of style sheets, hypermedia documents should encode the logical structure of the presentation while a separate hypermedia style sheet defines the mapping from the document structure to a final presentation format. Since the expressive power of current style sheet languages and their underlying models is insufficient to encode the temporal and structural characteristics of hypermedia documents, we propose a list of features which need to be included in such a language. The presentation environment model needs to include the aspects mentioned above: media specific characteristics, possible anchor visualizations, link behavior, and allowable temporal and spatial alignments.

CSS1 does not qualify as a hypermedia style sheet language because it does not support structural transformations. Additionally, since temporal structures are absent in HTML, these are not accounted for either. DSSSL and XSL are better candidates since they feature both a powerful query language and a language to express structural transformations on hierarchically structured documents. However, their two dimensional (page-based) presentation model cannot deal with layout schemes which are not based on the traditional text-flow model, and it is completely unable to deal with temporal scheduling. While both DSSSL and XSL provide hooks for extensions, developing such extensions for the complex spatio-temporal alignment of multimedia is difficult, especially because there exists no commonly accepted output model for hypermedia documents which is both sufficiently powerful to meet the requirements sketched above, and sufficiently high-level to avoid unnecessary complexity in the style sheet design process.

References

[1] Sharon Adler, Anders Berglund, James Clark, Istvan Cseri, Paul Grosso, Jonathan Marsh, Gavin Nicol, Jean Paoli, David Schach, Henry S. Thompson, and Chris Wilson. A Proposal for XSL, August 1997. W3C Note for discussion only. Available at http://www.w3.org/Submission/1997/13/Overview.html.

[2] Tim Bray and C. M. Sperberg-McQueen. Extensible Markup Language (XML): Part I. Syntax, March 1997. W3C Working Draft; Available at http://www.w3.org/pub/WWW/TR.

[3] International Organization for Standardization/International Electrotechnical Commission. Information technology — Processing languages — Document Style Semantics and Specification Language (DSSSL), 1996. International Standard ISO/IEC 10179:1996.

[4] Rohit Khare and Adam Rifkin. XML: A Door to Automated Web Applications. *IEEE Internet Computing*, 1(4), July - August 1997.

[5] Håkon W. Lie and Bert Bos. Cascading Style Sheets, level 1, December 1996. W3C Recommendation; Available at http://www.w3.org/pub/WWW/TR/REC-CSS1.

[6] Jacco van Ossenbruggen, Lynda Hardman, Lloyd Rutledge, and Anton Eliëns. Style Sheet Support for Hypermedia Documents. In *Hypertext'97 — The Eighth ACM Conference on Hypertext*, pages 216–217, April 1997.

A Simple Model for Adaptive Courseware Navigation

D. Pilar da Silva, R. Van Durm, E. Duval[1], H. Olivié
Departement Computerwetenschappen, Katholieke Universiteit Leuven, Leuven, Belgium
E-mail: {Denise.PilarDaSilva, Rafael.VanDurm, Erik.Duval, Olivie}@cs.kuleuven.ac.be

Abstract

This paper presents ongoing research and development on creating adaptive background material for a last year university course on 'multimedia modeling and programming', hereafter called *the course*. The course is organized around concepts, which are explained by documents. The documents and linking information are stored in a database. Concepts have explicit relationships with documents and with other concepts. Each document has an associated level of difficulty. The student is guided towards appropriate documents based on information about his knowledge of each concept.

1 Introduction

According to Brusilovsky [Brusilovsky 96] there are several goals to be achieved with adaptive navigation support techniques, though they are not clearly distinct: (1) global guidance, (2) local guidance, (3) local orientation support and (4) global orientation support. Such techniques comprehend: direct guidance; sorting, hiding and annotation of links and map adaptation. Most of the existing Adaptive Hypermedia (AH) systems use link hiding or link annotation in order to provide adaptive navigation support. Both are very efficient techniques.

In the following sections we shall present the structure of the course and how this structure contributes to our approach to provide adaptiveness.

2 The Course

A primary concern for educational hypertext is the definition of an appropriate structure so that a student can easily and naturally find the most relevant information depending on his/her needs.

Irrelevant information and links overload student memories and screen estate. In order to overcome this problem, it is possible to rely on information about particular users (represented in a user model) to adapt:

- the content (*adaptive presentation*), and/or
- the links to be presented to that user (*adaptive navigation support*) [Brusilovsky 96].

For the moment, our approach is mainly concerned with adaptive navigation support.

We rely on *typed nodes* (concepts and documents) and *typed and weighted links* to represent the structure of the course, which is organized around *concepts* explained by a set of *documents*.

Links between concepts represent semantic interrelationships. At present, we consider only two link types for the concept-concept relationship: *is_prereq_of* and *is_specialized_by*, but we plan to augment the model in the future with other link types such as *is_related_to*, *is_similar_to* and *contrasts_with*. Links between concepts and documents will also be typed though this facility is not implemented yet.

The *documents* are multimedia objects, such as text segments, static figures or interactive demonstrations. The URL of these documents is stored in a database. Each document has an associated *level of difficulty* with respect to the concept it belongs to, which varies from *0* to *99*, where a higher weight means '*more complex*'.

3 Student Model and Adaptive Navigation

The user model currently deals only with knowledge about each concept. We initialize student knowledge (or *level of expertise*) for a particular user as *0* for every concept [Calvi & De Bra 97] and update this value after the student has visited a document related to the concept. We are aware that this is a rather naive approach, but we have adopted it for the time being because the course is intended for a last year university audience, assumed to be homogeneous.

The level of expertise determines the documents available to a student. Basic concepts have no prerequisites and can be accessed by a new student. Acquiring these *basic concepts* enables the student to consult documents related to more advanced concepts.

The current knowledge *k* of a student *s* about a particular concept *c* can be described as *k(s,c)* with *0 £ k(s,c) £ 99.*

Associated to the *is_prereq_of* relationship between two concepts, there is a *threshold* t, that represents the minimal level of expertise a student must attain on the prerequisite in order to access the more advanced concept.

The *set of relevant concepts* RC_s for a student can be defined by the rule:

RC_s = { *c* | *basic(c)* *Ú* [$ c': (c' *is_prereq_of* c) Ù (*k(s,c)* ³ t(c', c))]}

This means that a student can access basic concepts, as well as concepts whose prerequisites he masters sufficiently well.

The documents accessible to a particular student are those that belong to the *set of relevant documents* (RD_s), defined as follows:

RD_s = { *d* | *explains(d,c)* Ù *c Î* RC_s Ù [*(k(s,c) - d) £ diff_level(d) £ (k(s,c) + d)*] }

Here, *d* represents a document and d is a constant. The above expression means that relevant documents are those that explain a relevant concept with an appropriate difficulty level (as defined by *d).*

When a student visits a document, his/her level of expertise is updated in the following way:

if d Î RD_s *Ù k(s,c) < diff_level*(d)
then *k(s,c) ¬ diff_level(d)*

The work described in [Signore 97] explores similar ideas in a non-educational application, where the user has more control (for instance with respect to the threshold value).

4 Current Status and Future Work

Currently, all information is stored in an Access database and dynamically translated into HTML. There are several approaches to do this, varying from the more traditional CGI approach to server-side includes and newer technologies like *servlets*. At the moment, we rely on the Active Server Page technology developed by Microsoft, but we intend to use a new approach we developed, where a Java based web server instantiates the necessary classes from a persistence layer [Hendrikx, Duval & Olivié 97]. The application will be tested soon with students at K.U.Leuven.

At the moment we are working on the user-interface, which, we believe, plays a crucial role in an adaptive hypermedia system, especially in an educational framework.

In the near future, we want to elaborate the user model by including some cognitive characteristics, which are relevant for learning processes, like for instance cognitive style [Höök 96, Wilkinson 97], or reasoning abilities. We will enrich the prototype with *stereotypes* that will also be part of the user model. Thus, a student will only see the links that are associated with the stereotype representing his/her profile.

Another enhancement we are currently implementing is the dynamic drawing of local overview diagrams or concept maps that show the immediate neighborhood in order to help minimize cognitive overhead.

Tests can be easily accommodated in the model as simple documents, making it possible to evaluate any previous knowledge a student could have about the subject in order to allow him/her to skip known concepts. We can, as well, use those tests to assess and update the knowledge a student actually acquired when following the course.

REFERENCES

[Brusilovsky 96] Brusilovsky, P., Methods and techniques of adaptive hypermedia. *User Modeling and User-Adapted Interaction Journal* Vol. 6, pp. 87-129, 1996.
[Calvi & De Bra 97] Calvi, L. & De Bra, P. *Improving the Usability of Hypertext Courseware through Adaptive Linking*. Proc. of the 8th ACM Conference on Hypertext. pp. 224-225. April 1997.
[Davis 95] Davis, H. To Embed or Not to Embed. *Communications of the ACM*, 38(8), pp. 108-9. August 1995.
[Hendrikx, Duval & Olivié 97] Hendrikx, K., Duval, E. & Olivié, H. *HOME - The Next Generation.* WebNet'97 World Conference. Toronto. Canada. Nov 1997. Accepted.
[Höök 96] Höök, K. *A Glass Box Approach to Adaptive Hypermedia.* Ph.D. Diss. Stockholm Univ. October 1996.
[Signore 97] Signore, O. et al. *Tailoring Web Pages to Users' Needs.* Proc. of the workshop "Adaptive Systems and User Modeling on the WWW" - 6th International. Conference on User Modeling. June 1997. http://zeus.gmd.de/UM97/proceedings.html
[Wilkinson 97] Wilkinson, S. *Document Structure, Individual Differences and the Learning*

Process. Poster at the 8th ACM Conference on Hypertext. April 1997.

ACKNOWLEDGEMENTS

The financial support of CNPq/Brazil for D. Pilar da Silva's scholarship is gratefully acknowledged. The Socrates program of the European Commission provided financial support for the elaboration of the courseware, in the framework of the EONT project (http://hyperg.softlab.ntua.gr/EONT).

[1] Postdoctoral Fellow of the Fund for Scientific Research - Flanders (Belgium) (F.W.O.)

Resource-limited information retrieval in Web-based environments

Daan Velthausz and Henk Eertink

Telematics Research Centre
P.O. Box 589
7500 AN Enschede, The Netherlands
{D.Velthausz, H.Eertink}@trc.nl

Introduction

Commercial usage of the Internet increases. It is therefore likely that information providers will increasingly charge end-users for the information they provide on such networks. Hence, the need will also arise for information retrieval procedures that takes such costs into account. However, currently little support exists for such *cost-effective* information retrieval in a networked environment with multiple information providers. In this paper we present a framework for resource-limited information retrieval that enables a user to search for relevant information given time and cost constraints, e.g. dealing with information needs like

> retrieve the five most relevant images containing white monkeys as fast as possible, but within 1 minute, for less than $4.

We focus in our work on the 'retrieval strategy'. In our terminology, this strategy is a functional entity that is responsible for handling the query, that makes decisions on which information providers are queried, and which objects (if any) are being retrieved. In this paper we present such a strategy for resource-limited information retrieval (see also Velthausz, Eertink, Verhoosel, & Schot 1997), and discuss how estimate information can be obtained for web based information environments prior to the search. We validate the strategy by means of a WWW-based prototype.

Basic concepts

Our framework is based on Russell and Wefald's metareasoning decision theory (Russell & Wefald 1991). This decision theory allows one to select between so-called 'internal actions' and 'external actions', based on a notion of maximum expected action utility (reflecting the degree of 'usefulness'). Internal actions are, in our usage of this theory, the search actions, whereas the external actions are the decisions to retrieve particular objects. The expected utility is based on the provider-estimated information for the (query dependent) relevance and involved cost of the actions. These estimates need to be available prior to the execution of the best action. As the search proceeds, the system explores unknown parts of the information environment and updates its view accordingly. When relevant information has been found, the system must decide whether to retrieve it or continue searching trying to find 'better' information (e.g. more relevant or cheaper). Because the cost specified in the query refers to the total cost, the system continuously needs to decide how much budget should be spent on searching and how much on retrieving. This is the fundamental problem that we tackle in our resource-limited retrieval strategy.

Strategy

The user can specify the 'optimal' selection of information objects that fulfils the constraints, for example maximise the sum of the relevance of the retrieved results at minimum cost. The closer the results of the search are to the 'optimum', the higher the utility of the search session will be. It may be clear that when the search proceeds the current session utility is likely to change because cost and time are continuously being spent.

The outline of the strategy is:

1. Determine all possible start actions **A**, given the query.
2. Remove all actions which do not satisfy the time, money and/or relevance constraints.
3. When no actions are left, terminate the search session.
4. For each new action A*i* estimate its utility EU(A*i*).
5. Select the action Aα with the highest estimated utility.
6. Determine whether the estimated utility EU(Aα) is higher than the session utility U(**A**). If EU(Aα) is not higher than U(**A**) then terminate the search session, else carry out action Aα.
7. Update the strategy's view of the information environment, the time-spent, money-spent, **A**, U(**A**) and determine all new possible search and retrieval actions.
8. Go to 2.

In step 2 the strategy removes all actions that do not satisfy the constraints specified in the query. This means that for each action that is not removed, it is estimated that it can be executed within the remaining time and money budget. The cost estimates need not necessarily be single valued (e.g. the average) but may consist of multiple values, like minimum, maximum etc., or in general be a probability distribution function. By supporting different strategy variants the risk of overspending can be controlled. The variants differ in action filtering, action selection and session termination.

In step 4 and 5 the estimations for individual (search and retrieval) actions are computed and compared. This action utility should only be based on the results and cost related to the particular action. This means that the user formulated 'session utility function' should be adapted to reflect the cost and relevance aspects of the individual actions. To use the relation between the *estimated action utility* and *session utility* as a stop criteria, as is done in our strategy, the following criterion should be met for each estimated action utility function: *when the estimated utility of an action is higher than the current session utility and assuming perfect estimates, execution of this action should increase the session utility.*

Because searching is in general only useful to a user when (a property of) a relevant information object is retrieved, not only the cost of a *search* action itself should satisfy the constraints but also the (estimated) remaining search cost to find a relevant information

object as well as the (estimated) retrieval cost of this object. For example, suppose we have an session utility function: (sum of relevance)/cost. Then the estimated utility of a search action j will thus be E[retrieval cost] + E[relevance object j]/(E[search cost]. And the estimated utility of retrieval action j is E[relevance object j]/E[retrieval cost]. We show that these action utility functions meet the above defined criterion.
Let **A** denote a search session: a sequence of actions (A1 ... Ai); Aj is action j. Aj can also be denoted as Sj when Aj is a search action and Rj when Aj is a retrieval action. EU(Aj) is the estimated utility of action Aj; U(**A**) is the session utility and U(**A**.Aj) is the session utility after executing action Aj. It can be proven that when EU(Rj) > U(**A**) that U(**A**.Rj) > U(**A**) when relevance and cost estimates satisfy the above defined conditions. In our example U(**A**) = r/c (denoting session-relevance/session-cost) and EU(Rj) = rj/cj (denoting relevance of action Rj /cost of action Rj) then U(**A**.Rj) = (r+rj)/(c+cj). When rj/cj > r/c this implies that (r+rj)/(c+cj) > r/c. When Rj is the retrieval action included in the estimation of U(Sj) and perfect relevance and cost estimations are available, then U(Rj) > U(Sj), and thus U(Sj) > U(**A**). Hence U(**A**.Sj.Rj) > U(**A**).

When the best chosen action is carried out, the current view of the information environment is updated accordingly and the strategy continues the search session.

Estimation of cost and relevance

It is the responsibility of the information provider to provide the (query dependent) estimates to the strategy. This is vitally important for the effectiveness (in terms of recall and precision) of the strategy. To get an idea how the estimates can be obtained, we briefly sketch the ideas behind the mechanisms of the ADMIRE information model (Velthausz, Bal & Eertink 1996), that we use in our IR system. The ADMIRE model provides a general framework for modelling any information, regardless of type, size or abstraction level, including web-based information environments. Via a notion of different kinds of information objects and (informed) relationships, it facilitates a gradually expanding action search graph, with labelled transitions that can be used to estimate the (query dependent) relevance and involved cost of the actions. The problem is that the relevance of information objects is query dependent and can only be determined when either the object is accessed or fully indexed. Given the wide range of possible queries in a large and dynamic environment, it is impractical to reuse relevance estimations of previous searches nor fully index all objects. The ADMIRE information model facilitates aggregation and propagation of information that characterises reachable information objects. The composite relationships in the object hierarchy enable a bottom-up propagation and aggregation of the lower layered object characterisations. This information can subsequently be used to estimate the relevance of unexplored information objects. This use of summarised information to describe particular aspects of the lower layered nodes in a hierarchy, has also been reported in (Garcia-Molina, Gravano & Shivakumr 1996) for the content-characterisation for (hierarchical) databases containing textual documents. We have adapted some of their ideas and included them in our prototype for web-based information to provide vector based keyword text-retrieval. We use a standard term frequency weight scheme suitable for dynamic collections with varied vocabulary conform experimental evidence of (Salton, G., Buckley C., 1988).

In addition to estimating the relevance, an accurate estimation of the remaining search cost to find the requested information objects as well as the retrieval cost of these objects is needed. As the cost may (partly) depend on the query, the structure of the query (e.g. the operators, the requested number of objects etc.), can be used to estimate the remaining search and retrieval cost. Although the ADMIRE model has not been developed to support cost estimations, it can simply be extended using the bottom-up propagation and aggregation of cost-related information. Instead of the actual distribution function, derived measures such as average, minimum, maximum, variance, etc. can be used.
A potential problem with upwards propagation of information is the topicality of the aggregated information. Since this information is obtained prior to the actual search, changes in costs, adding new or deleting information objects might not be directly reflected in the aggregated information. Furthermore, a consequence of the use of the aggregation methods is that the cost and relevance estimates are independent of each other and therefore may lead to inaccurate utility estimations.

Using ADMIRE on the Internet

Although the ADMIRE information model was not developed for web based information environments we want to determine whether the ideas can be applied in web based information environments. As web based information environments are typically not hierarchical and are likely to contain (many) cycles, we need to adapt the aggregation and propagation process.
A quick view on the world wide web shows that it appears to have hierarchical (sub)structures incorporated by the information providers, e.g. home page seen as a root for a site. Based on this observation we assume that the authority performing the aggregation process, assigns root-page(s) or -site(s) from which a 'pseudo' hierarchical structure can be derived. We have looked at different ways to deal with cycles. Dealing with cycles effectively implies cycle (membership) detection during the aggregation process. A simple approach is to skip links in the aggregation process that introduces a cycle. In this manner a directed a-cyclic graph is obtained and the standard aggregation process can be applied. The drawback of this method is that the decision of which links to skip, which is determined by the cycle avoidance algorithm, has effect on the aggregation results. Another better but more time consuming approach is to construct a directed a-cyclic graph for each node that is part of a cycle, and apply the standard aggregation process. Another approach is to consider nodes that are part of a cycle as a single 'virtual node' during the aggregation process. This results in one aggregation description for all objects of the cycle. A drawback of the latter approach is it may reduce the information structure to a few or even a single node when many cycles exists.
A different approach that loosens hierarchical relationships to reachability relationships (object a is reachable from object b, either directly or via intermediate objects) is based on dynamic routing algorithms. As the links in a web based information environment are unidirectional (from referring object to the referred object), pushing (or flooding) information from the referred object to the referring object(s) is not possible as the referred object does in general not know the identity of his referring object(s). This means that a

kind of pull principle needs to be applied in the aggregation process. In the full paper we describe the whole aggregation process for web-based information environments in detail.

Evaluation

In our prototype we tested the algorithm on our own intranet to test the principles behind the mechanism. In practice it may be better (and much faster) to index a particular web site (by means of inverted file (keyword to url)) and use this index information in the aggregation process for pages that link to this website, e.g. a broker referring to n different websites. When the individual pages are considered to represent indexed websites we can test the ideas behind the mechanism in a relatively small and stable environment, our intranet web-pages.

References

Garcia-Molina H., Gravano L., Shivakumr N., 1996. Finding Document Copies Across Multiple Databases, In *Proceedings of 4th Int. Conference on Parallel and distributed Information Systems,* PDIS'96.

Russel S., Wefald E. *Do the Right Thing,* MIT Press, 1991

Salton, G., Buckley C., Term weighting approaches in automatic text retrieval. In *Information Processing and Management,* Vol. 24, pp 513-523, 1988

Velthausz, D.D., Bal, C.M.R., Eertink, E.H. A Multimedia Information Object Model for Information Disclosure. In *Proceedings of the MMM'96, Third International Conference on MultiMedia Modelling,* Toulouse, France, pp 289-304, 1996.

Velthausz, D.D., Eertink, E.H., Verhoosel, J.P.C. Schot, J. Resource-limited information retrieval in a distributed environment. In *Proceedings of AAAI-97 Worksop Online Search,* Providence, USA, pp 114-121, July 28, 1997

Bestemd voor conferentie Werkgemeenschap Informatiewetenschap, donderdag 27 november 1997, "Let your browser do the walking". Extended abstract, November 3, 1997.

Filter Browsing on the World Wide Web

Jeroen Vendrig, Marcel Worring, Arnold Smeulders

{vendrig, worring, smeulder}@wins.uva.nl, Intelligente Sensorische Informatie Systemen, Universiteit van Amsterdam, Faculteit WINS, Kruislaan 403, 1098 SJ Amsterdam.

Abstract

Navigation through large multimedia information spaces depends on automated indexing of the multimedia data and dynamic interpretation of the user?s information needs. In full-text retrieval and especially in image retrieval, research has focused on similarity ranking techniques to dynamically respond to user queries. However, similarity ranking can lead the user to a homogeneous collection of undesired data. In this article we describe the Filter Browsing approach that is based on differences, rather than similarities, between information entities. Users can navigate through large databases by zooming in on parts of the database that suit their information needs. By iteratively limiting the information space, the user is able to end up with a small amount of relevant information items that can be inspected one by one. The application of Filter Browsing in the ImageRETRO system is described. For evaluation of ImageRETRO a pictorial database was filled with 10,000 images acquired automatically from the World Wide Web.

Introduction

In the past decades large archives of digital textual documents have been created. Due to a decrease in cost of acquiring and storing pictorial data, image- and videodatabases gain popularity as well. Finding information entities in such databases requires additional retrieval techniques. Indexing of data by adding descriptions to the items is intuitively the best way to assist users. However, the generation of descriptions has some disadvantages. In large data collections, manual indexation would be very labor intensive and thus costly. Besides, humans will assign descriptions to information items based on their own point of view. With regard to images, in [JAI96] it is concluded that "keywords often provide a better description of the person who assigns them than they do the image".

Therefore, the assignment of keywords and descriptions should be automated. For full text documents elaborate indexing techniques are available, reducing large chunks of text to a short list of relevant words. Extracting textual descriptions from images is however less successful. The translation from an image object to keywords requires a great deal of knowledge about the domain. Further, unlike words in a text, objects in an image can?t be distinguished easily.

The same problem holds for Query by Association, the information retrieval method in which user input is a choice of one of the navigation controls, e.g. multimedia hypertext. Since the controls have to be based on the image content, the user-friendly approach of Query by Association has the similar disadvantages as keyword queries.

Context can be useful to reduce a multimedia retrieval problem to a full text retrieval problem. In [SRI95] the Piction system is described that identifies human faces in newspaper photographs using visual semantics and natural language processing. Piction associates terms in the caption of an image with a photograph, or even a part of the photograph. The indexing is performed fully automatically by a computer system. The retrieval of the images is based purely on descriptions.

Since context is not always available and indexing on information encoded in the image itself would be less speculative, multimedia research focuses on indexation on image features. Image features are non-information-bearing [GRO94] metadata derived directly from the pictorial data. Query by Pictorial Example is currently the most used method to retrieve images.

In the next section the Query by Pictorial Example method will be explained in detail. Next, it is shown how disadvantages of the method can be avoided when combined with concepts from Query by Navigation. The resulting method is called Filter Image Browsing. Results of an experiment with a Filter Image Browsing implementation will be described, followed by conclusions and suggestions for general application of Filter Browsing in multimedia information retrieval.

Query by Pictorial Example

In 1980 [CHA80] introduced the concept of Query by Pictorial Example (QPE) which was based on the Query By Example (QBE) concept for alphanumeric data. Query by Pictorial Example is the retrieval method in which the query for images itself consists of one or more images. The result of the query is a similarity ranking of a set of images according to the similarity in value for some image features. Thus QPE is not entirely comparable to QBE. In QBE a user formulates a query by filling in example values and conditions in fields in a skeleton table. The result of QBE is a set of records that exactly match the given conditions. QPE caused a shift in image retrieval research that is comparable to a shift in the research field of full-text retrieval described in [HUI96].

Although the features for computing the similarity differ a lot, Query by Pictorial Example uses the same principle as full-text information retrieval to satisfy a user's information need. A user can for example provide the image retrieval system with a picture of a red car, asking for all pictures that look alike. The system computes the values for the features of the example image and compares them to the values of the images in the database. In Query by Pictorial Example, the system's purpose has changed from giving deterministic answers to minimizing the user's search and evaluation efforts.

Besides image features and similarity ranking techniques, in QPE the example images given by users is of great importance. We distinguish three types of Query by Pictorial Example:

- Query by external Pictorial Example: the image consists of sensoric data sampled from the real world;
- Query by internal Pictorial Example: the user selects an image that is present in the database;
- Query by Construction (also known as Query by Canvas or Query by Sketch): the user creates a new synthetic image for input. The constructed image can contain parts of real images.

For systems that desire quick user interaction Query by internal Pictorial Example is the most suitable. Query by external Pictorial Example requires the user to have images beforehand, which is only the case in specific applications. Query by Construction requires a lot of artistic effort from the user. In Query by internal Pictorial Example the user can simply tell the system which images are most similar to the images he desires to find.

Query by internal Pictorial Example has two major disadvantages. First, the system has to provide the user with images from which the user can choose an example image. The search for a suitable example image can be time consuming. Secondly, although the Query by internal Pictorial Example method can be applied iteratively, the use of similarity ranking can cause premature convergence. Premature convergence can occur when the top ranked images are presented to the user as example images for a query refinement. The purpose of the system is to present a set of images that all look alike to the user. We call the convergence premature when the set of images does not yet comply with the user?s information need, but it is not useful anymore for a new query, since the result will be the same.

Filter Image Browsing

In the ImageRETRO system (Image RETrieval by Reduction and Overview) [VEN97], Filter Image Browsing is introduced as a new retrieval method in which a user iteratively selects dynamically generated clusters of images. Filter Image Browsing is based on Query by internal Pictorial Example, but uses Query by Association concepts to overcome the disadvantages of the former method.

Filter Image Browsing uses the concept of overview that is common for Query by Association systems. The system shows images that are representatives for parts of the database. The user then selects a representative image and thus zooms in on a part of the database. Next, this process of overview and reduction is repeated. The overview concept and the gradual specification of the user query solve the two major problems for Query by internal Pictorial Example. However, Query by Association heavily depends on manual classification of the information entities and groups and is thus not feasible.

Filter Image Browsing can be seen as an overlay over Query by internal Example, or as the addition of a dynamic zoom function for image databases to Query by Association. The Filter Image Browsing system first presents an overview of images that are considered representative for parts of the image collection. The user selects an example image from the overview after which a similarity ranking is performed. A large percentage (typically 25%) of the top ranked images forms the new subset of images. The cycle of overview, image choice and filtering is repeated until the user decides to end the retrieval session.

Thus Filter Image Browsing can use similarity ranking, where premature convergence is prevented by applying the overview techniques to the entire subset of images instead of just the top ranked images. In the next section, Filter Image Browsing will be described in more detail. Further, some side effects of Filter Image Browsing will be discussed.

Overview and reduction

Filter Image Browsing is based on two functions, viz. overview and reduction. The overview function has to provide the user with a small yet representative set of images, so that she can choose a good example image. The reduction function has to make a distinction between images that are likely to be desired and images that are not, based on the information (the selected example images) available.

The use of reduction and similarity ranking does influence the overview function, since all images in the database must be accessible by choosing at least one of the images presented to the user. Thus the overview images have to cover the entire set of images in the selected part (also called cluster) of the database. Therefore, we introduce the cover constraint stating that every image in the current set must be part of at least one of the possible new clusters after a reduction.

Side effects

Two positive side effects of Filter Image Browsing are worth mentioning. Both side effects are caused by the variation in active image sets during a retrieval session.

First, the use of smaller sets of images during the session results in an increase of processing speed. IBM?s QBIC already applies filtering techniques for the purpose of speeding up computations solely [FLI95]. Thus the system can use computationally expensive features once the size of a reduced image set allows it.

Secondly, the system can decide to use certain features when the image set is suitable for it. Features that are valuable to make a general distinction between image clusters in the beginning of the retrieval session, may loose their value after the distinction has been made. On the other hand, some features can be applied only when certain preconditions are met. Further, statistics can be used to determine the discriminatory power of each feature in a state of the retrieval process.

A negative side effect of Filter Image Browsing is the possible loss of relevant images after a reduction. However, since those images were not ranked highly relative to the seed image, the lost images would be hard to find in any case. The user can minimize the risk of loosing images by changing the reduction factor, at the expense of a longer retrieval session.

Evaluation

We have evaluated the ImageRETRO system on a database containing 10,000 images. The images come from a large and diverse domain, the World Wide Web. A robot has gathered the images by following hyperlinks and could easily be used to extend the number of images in the database. For every image index values were computed for 9 simple image features, primarily based on colors.

The ImageRETRO system will be available via the World Wide Web in November.

Since a large scale and unambiguous evaluation of the system by human users was not feasible, a simulation was set up. The goal of the simulation is to generate results that make it possible to compute two evaluation criteria. The first criterion tests the average decrease in precision compared to the reduction of the image sets. The decrease in precision has to be lower than the decrease in size of the image set.

The second criterion compares the ImageRETRO system to a straightforward implementation of the Query by internal Pictorial Example method. For ImageRETRO to be successful, the number of images found after a reasonable amount of choices must be higher than the same figures for Query by internal Pictorial Example.

The simulation results are satisfying. Although Filter Image Browsing does cause the loss of desired images, it performs better than Query by internal Pictorial Example.

Conclusions and future research

Filter Image Browsing combines two methods that have proven to be successful: Query by internal Pictorial Example from the image retrieval field, and Query by Association from hypertext research. Filter Image Browsing assists the user in zooming in on an image database, thus specifying his information need iteratively. Dynamic clustering of image groups is the most important feature of Filter Image Browsing. It was shown that the cover constraint is necessary to guarantee access to every image in the database.

In future research, we will apply a similar approach to full-text databases, so that users can find documents by iteratively choosing automatically indexed keywords from the database. In cooperation with Origin/Medialab a website for Oogziekenhuis (Eyehospital) Rotterdam is currently under development using such techniques. Later, combination of image and full-text Filter Browsing will lead to a uniform approach for multimedia information retrieval.

Further, we plan to extend Filter Browsing systems with innovative interfaces, such as Medialab?s Aquabrowser [VEL97]. The combination of user-friendly interfaces and powerful adaptive retrieval systems has to result in intuitive access to multimedia databases.

References

[CHA80] N.S. Chang and K.S. Fu. Query-by-pictorial example. IEEE Transactions on Software Engineering, SE-6(6):519-524, 1980.

[GRO94] W.I. Grosky. Multimedia Information Systems. IEEE Multimedia, 1(1):12-24, 1994.

[FLI95] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele and P. Yanker. Query by Image and Video Content: The QBIC System. IEEE Computer, 28(9):23-32, 1995. (demo)

[HUI96] T.W.C. Huibers. An Axiomatic Theory for Information Retrieval. PhD Thesis, Universiteit Utrecht, The Netherlands, 1996.

[JAI96] R. Jain. Infoscopes: multimedia information systems. In Multimedia Systems and Techniques, B. Furht, ed., Kluwer Academic Publishers, Boston, 1996, pp. 217-253.

[SRI95] R.K. Srihari. Automatic Indexing and Content-Based Retrieval of Captioned Images. IEEE Computer, 28(9):49-56, 1995.

[VEL97] W.A. Veling, Aqua Browser - The answers, Origin/Medialab report AQUA-001. http://www.medialab.nl/research/aqua/whatisit.htm

[VEN97] J. Vendrig , Filter Image Browsing: a study to image retrieval in large pictorial databases. Msc Thesis, University of Amsterdam, The Netherlands, February 1997. http://carol.wins.uva.nl/~vendrig/thesis/

Ontsluiting van Web sites naar analogie met de inhoudsopgave van een boek: voordelen en toepassing.

Nederlands Instituut voor Wetenschappelijke Informatiediensten
(NIWI: www.niwi.knaw.nl)

R.E. de Vries en N.J.I. Mars

Inhoudelijke ontsluiting van Web sites volgt grofweg drie benaderingen:

- bij weinig aangeboden informatie wordt volstaan met een overzicht op de home page,
- bij een veelheid aan informatie of een complexe struktuur wordt ofwel een zoekfaciliteit geboden op alle inhoud van de site,
- of er is een handmatig aangelegde alfabetische lijst met termen die ontsluiten.

De laatste mogelijkheid komt relatief weinig voor, terwijl deze een gebruiker van een Web site veel meer steun biedt bij het bepalen en vinden van gewenst informatie- of diensten aanbod. De variant waarbij uit alle materiaal op een site geautomatiseerd een index gegenereerd wordt, is een mengvorm die weinig verschilt van de zoekbenadering omdat de termen niet inhoudelijk beredeneerd zijn maar bestaan uit de woord- en begripskeus in het Web aanbod zelf.

De drie benaderingen worden met voorbeelden op de volgende categorieën vergeleken:

- de toegevoegde waarde ten behoeve van navigatie om een Web site bezoeker adekwaat aanbod te laten bepalen en lokaliseren
- de mate waarin het een verdwaalde Web site gebruiker een uitweg en alternatief biedt
- de investering die realisatie vraagt bij Web site aanleg en onderhoud

De home page benadering kent geen aparte inhoudelijke ontsluiting, dus geen toevoeging aan navigatie. Er is geen alternatief bij verdwaald zijn, evenmin vergt realisatie extra inspanning.

De zoekfaciliteit benadering biedt een alternatief voor navigatie maar laat het aan de gebruiker over om met herhaald proberen een idee te krijgen van het aanbod. Om dezelfde reden biedt deze faciliteit weinig wegwijzer hulp. Tot stand brenging is eenvoudig.

De benadering van een handmatig aangelegde alfabetische lijst met termen is eveneens een alternatief voor navigatie maar geeft wel direkt een indruk van het aanbod. Een dergelijke lijst kan bovendien door synoniemen of alternatieve benamingen rekening houden met eigen terminologie en verwacht woordgebruik door onderscheiden groepen gebruikers.
Ook bij verdwaald zijn is dit een voordeel. Het navigeren naar een gewenste Web pagina, volgt direct uit aanklikken in de lijst.
Aanmaak en onderhoud vergen vergelijkenderwijs bij een handmatige lijst de meeste investering.

De mengvorm van een geautomatiseerd aangemaakte lijst van woorden als benadering, is eenvoudiger te realiseren maar biedt weinig wegwijzer hulp doordat de termen dezelfde blijven als al op de Web pagina's in gebruik zijn en daar ook al voor moeten komen. Navigatie vanuit een lijst is op zich gemakkelijk maar biedt zo toch weinig alternatief.

Vervolgens wordt de Web site van het Nederlands Instituut voor Wetenschappelijke Informatiediensten (NIWI) besproken als case van ontsluiting met een lijst met termen die vooraf inhoudelijk gewogen zijn, naar analogie van de inhoudsopgave van een boek.

Aangegeven zal worden:

- hoe de termen inhoudelijk tot stand komen,
- hoe de relevantie van de termen voor de verschillende groepen Web site gebruikers, nagegaan en bewaakt wordt,
- hoe de inhoudsopgave met termen praktisch aangelegd en onderhouden wordt, gegeven een Web site die aan verandering onderhevig is

Het inhoudelijk bepalen van de termen steunt op analyse van het aanbod van NIWI informatiediensten zoals dat op specifieke Web pagina's vorm krijgt: die termen zijn geschikt die het aanbod specifiek genoeg verwoorden en (daarmee) vervolgens ook uniek naar een bepaalde NIWI Web pagina kunnen verwijzen. Synoniemen en het geven van meerdere ingangen aan samengestelde termen, verhogen de waarde van de lijst onder behoud van die unieke pagina referentie.

> NIWI onderhoudt bijvoorbeeld databanken met bibliografische bestanden. Deze zijn er voor verschillende wetenschappelijke disciplines. In de inhoudsopgave voor de Web site zijn daarom hoofdtermen als "bibliografische verwijzingen" maar ook "tijdschrift artikelen" of "auteurs" (en met de discipline als subterm) opgenomen die direkt verwijzen naar de pagina waar in de bestanden gezocht kan worden. Omringende pagina's in het Web design hebben geen verwijzing omdat deze alleen functioneel zijn voor stapsgewijze navigatie vanaf de home page.

De bepaling van de relevantie van termen voor verschillende groepen Web site gebruikers gebeurt door bijvoorbeeld te vergelijken met woordkeus en begrippengebruik in door NIWI bestreken wetenschapsgebieden.

> Een voorbeeld zijn de data-archieven van NIWI. Gebruikelijke termen in publicaties met kwantitatieve analyses zijn: "databestanden", "secundaire analyse" en het gebruik van "data beschikbaar stellen". In de Inhoudsopgave verwijzen de eerste twee termen naar de Web pagina waar de catalogus doorzocht kan worden, de derde term naar een registratiefaciliteit voor databestanden.

De praktische aanleg en onderhoud van de inhoudsopgave beperkt zich doordat alleen die Web pagina's verwijzingen hoeven te krijgen waar een (informatie)dienst of aanbod plaats vindt. Tussenstappen en dergelijke in het Web design, komen daarmee niet in aanmerking. Nederlandse en Engelse termen worden in een plat ASCII bestand bijgehouden, telkens met de bestandsnaam van de corresponderende Web pagina. Een Perl programma genereert hieruit de "Table of Contents" en "Inhoudsopgave" voor de Web site en brengt de termen aan in de betrokken pagina's zelf volgens HTML MetaTag syntax. Het laatste anticipeert het indexeren van NIWI Web aanbod door robots en search engines. Woordenboeken en synoniemenlijsten worden vanaf het Internet gebruikt, wat het deels automatiseren van onderhoud zal vergemakkelijken.

Conclusie

De door NIWI gekozen benadering voor ontsluiting van de Web site, is die van een handmatig aangelegde alfabetische lijst naar het model van de inhoudsopgave van een boek. Termen kunnen synoniemen hebben en onder meerdere ingangen voorkomen maar verwijzen telkens naar één unieke Web pagina. Voor gebruikers die verdwaald zijn of een zo direkt mogelijke navigatie wensen, heeft dit grote voordelen. Door bij toekenning van termen te redeneren vanuit het dienstenaanbod van NIWI en niet vanuit elke individuele Web pagina, is het aantal verwijzingen relatief beperkt. Ook de nadere criteria van specifiekheid en uniek refereren, houden de lijst met termen beperkt. Daardoor kunnen de voordelen voor de gebruiker van deze wijze van Web site ontsluiting, samengaan met beperkte investeringen aan onderhoud.

An Electronic Commerce Paradigm for Agent-Based Information Discovery

B.C.M. Wondergem, P. van Bommel, T.W.C. Huibers, and Th.P. van der Weide
Computing Science Institute
University of Nijmegen
Toernooiveld 1
NL-6525 ED, Nijmegen
The Netherlands
bernd@cs.kun.nl
http://www.cs.kun.nl/~bernd/

October 22, 1997

Keywords: Information Discovery, Intelligent Agents, Negotiation, Information Retrieval and Filtering, Electronic Commerce, Multi-Agent Systems.

Accepted at the Interdisciplinaire Conferentie Informatiewetenschap 1997.

Abstract

This article proposes an electronic commerce paradigm for agent-based Information Discovery (ID). ID is the synthesis of Information Retrieval and Information Filtering, thus coping with dynamics from both fields. Agent technology, having a number of valuable properties to attack problems with dynamics, is used in ID, resulting in a paradigm with three types of agents: user, broker, and source agents. Cooperation, which is the keyword for the agents to fullfil ID tasks, is seen as negotiation, thus obtaining a very general and flexible way to describe agents' interaction. The possibilities for ID agents to negotiate are investigated. In addition, remarks stemming from the ID background are made about other issues concerning negotiation such as protocols, strategies, and coordination.

1 Introduction to Agent-Based Information Discovery

Information Discovery (ID) (see [WBHW97b]) is the synthesis of Information Retrieval (IR) and Information Filtering (IF). For a comparison of IR and IF see [BC92]. In ID, we consider a population of users having dynamic information needs. These information needs are to be satisfied with information out of dynamic information sources. To facilitate this process, information brokers act as intermediaries between users and sources. Brokers compare the wishes of users with the available information in the sources.

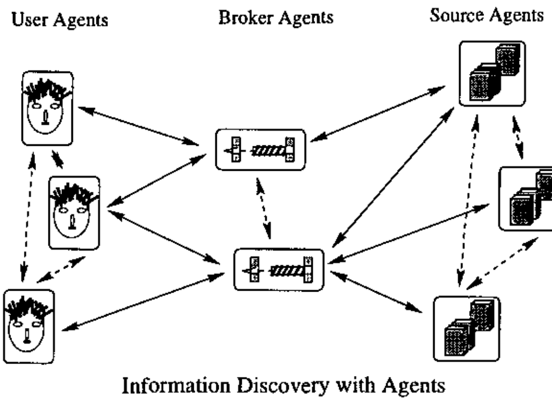


Figure 1: Conceptual ID Paradigm

Figure 1 sketches the ID paradigm at a conceptual level (as proposed in [WBHW97a]). Note that it is not possible for users to communicate directly with the information sources. Since agent technology (see e.g. [WJ95]) fits well in the ID paradigm (especially communication, intelligence, and proactiveness), the entities of the paradigm are seen as agents: user agents, broker agents, and source agents, respectively. The 3 collections of agents are called *spaces*.

Figure 2 sketches a more implementation oriented form of the ID paradigm. The agents reside in a networked world, made up of sites reachable by links, where on each site communication can take place with the ID constraint. This is one of many possible ways to translate the conceptual paradigm to a more concrete paradigm.

The arcs in Figure 1 visualize possible communication in ID: user–broker, broker–source, and

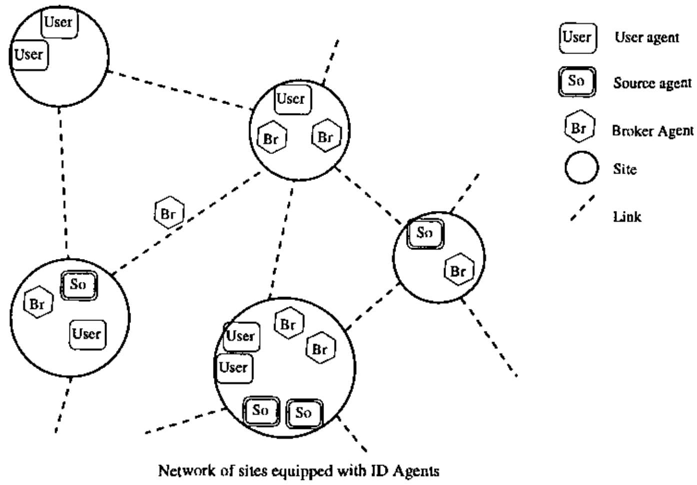


Figure 2: Implementation Oriented ID Paradigm

same type (intra-space) communication. Through communication, the agents effectuate their dynamic cooperation. A general and flexible manner to implement cooperation in a dynamic environment is *negotiation*.

To effect negotiation, a basic assumption is that the services of agents are not for free and that their costs are negotiable (within limits). This conforms to the modern trend on the Internet that the information available in the sources involves costs. The costs of information can be defined in terms of access-costs, transportation costs, complexity costs, etc. In addition, broker agents will also charge their clients. By negotiating, the agents in the ID paradigm form a *dynamic economic market*. For another multi-agent ID application see [Mar97].

The goal of this article is to show how electronic commerce can serve as the basis of an Information Discovery paradigm when negotiating agents are used.

Section 2 investigates the possibilities for negotiation amongst agents in the ID paradigm. Section 3 touches upon a number of aspects of negotiation, looked upon from an ID perspective. Section 4 offers concluding remarks and directions for further research.

2 Possibilities for Negotiation in Information Discovery

Basically, there are 2 types of negotiation in the ID paradigm: inter-space (among agents of different types) and intra-space (among agents of the same type) negotiation:

1. **Inter-space** negotiation among users and brokers or among brokers and sources involves the price/quality ratio of information (basic IR and IF).

2. **Intra-space** negotiation involves team formation and intra-team negotiations.

Figure 3 provides a scheme to characterise the types of negotiation. The circles represent the types of agents, and the arrows represent messages sent. Several types of negotiation will be described below. Starting at a certain agent type, negotiation follows the direction of the arcs, i.e. only outgoing arcs can be followed. In every type of negotiation between of agents, 1-1, 1-N, M-1, and M-N forms are possible.

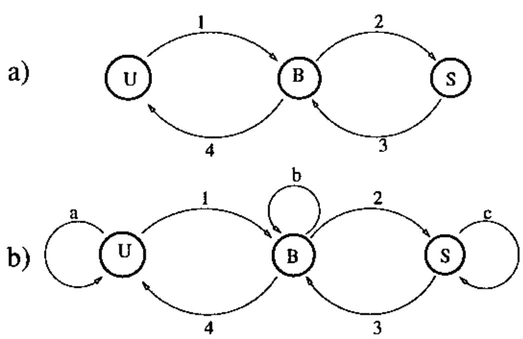


Figure 3: Communication Possibilities

2.1 IR and IF as Inter-space Negotiation

The ID paradigm is a synthesis of IR and IF. Both aspects of ID are covered in inter-space negotiation.

In figure 3a), a 1234-cycle represents the IR related task of query processing. A user fires a query, which contains a price/quality ratio, to a number of brokers (1). These brokers pass the query on to a number of sources to inform about a an offer (2). The offers (3) are sent as bids to the user (4). The user then selects *only* the most favourable bid (if not too expensive).

A 3412-cycle in figure 3a) represents IF. A source offers a number of documents to a number of brokers (3). The brokers pass the offer on to a number of users (4) which inform the broker about their bid (1). *Every* interested user is considered, to maximize payoff. The broker agent informs the source about the process (2).

Without negotiation, the above plans of attack collide with the standard notions of distributed IR and IF (viz. non negotiating brokers such as Lycos[1] and AltaVista[2]).

Negotiation can be effectuated by cycles of communication. For instance, a (14)*-cycle, where $*$ denotes iteration, (user-broker) can represent negotiation about the exact user query or about the broker's fee. The best bid after round n serves as the price to beat for round $n+1$. The negotiation terminates if no agent can beat the best price so far. As another example, consider 1(23)*4-cycles, in which the negotiation between broker and source is an iterative process. A (1234)*-cycle represents an iterative form of the standard IR negotiation.

2.2 Teams of Agents: Intra-space Negotiation

When an agent is not capable of achieving its goals on its own, the agent tries to form a team and achieve the (common) goal as a group. The team acts as a *single party* in the negotiations. The two aspects of teams, *team formation* and *intra-team negotiation*, involve the reflexive links in figure 3b) (links a, b, and c).

Reasons for ID agents to form a team:

1. **User agents** can form a team based on their queries or user profiles to place a common request at a number of brokers and pay only once.

2. **Broker agents** can make aboutness decisions cooperatively to obtain better efficiency

[1] Available from: http://www-nl.lycos.com/
[2] Available from: http://www.altavista.digital.com/

(filter teams) or to obtain better quality (using a structural hierarchy of composed brokers). In a non-negotiating way, MetaSearch[3] and MetaCrawler[4] are examples of the latter.

3. **Source agents** can cooperate if the information they carry is complementary or related, for instance. This can result in a structure not unlike a hyperindex (see e.g. [Bru93]).

Team formation is initiated by an agent incapable of achieving its goal or not making enough profit on its own. The agent has to search for agents with similar goals, so as to form a common goal which is supported by all the members of the team. In principle, a very large number of teams can be formated. The negotiations, from which agents can step down at any moment, has to establish a suitable team efficiently. When a team is formated, a team coordinator can be assigned to facilitate coordination within the team.

When a team of agents negotiates as a single party, the decisions of the team have to be supported by all the team members. For example, each agent in the team must be satisfied with its share of the total fee. Intra-team negotiations have to take place to establish this. The agents should take into account that the team can break up if one of the agents receives a better offer during the negotiations.

2.3 Other Negotiations in ID

Next to the basic IR and IF functions, possibly with teams, other forms of negotiation are possible in the ID paradigm. In all cases, all the parties involved should be better of in the end.

For instance, user agents can cooperatively expand their queries and user profiles. In addition, source agents can send out sample documents as advertisement material. Also, relevance feedback can be seen as a negotiation process. Merging of documents can also be viewed as a negotiation process. Next, agents can communicate about their domain knowledge. It is also possible for agents to form long-term contracts by forming coalitions. For example, a broker agent with specific domain knowledge can form a coalition with a source containing information about that specific topic.

[3]Available from: http://metasearch.com/

[4]Available from: http://www.metacrawler.com/

3 Other Aspects of Negotiation

This section touches upon a number of aspects of negotiation, seen from an ID perspective: coordination strategies, protocols, negotiation strategies, negotiation mechanisms, and properties of ID agents.

Conditions stemming from the ID context are: dynamics of agents influence population during negotiations (open environment), bounded rationality (due to e.g. limited computation time or reasoning capabilities about the enviroment), and privacy (obvious for users, also concerning competition). The negotiation strategies should be able to cope with these problems.

An important aspect of negotiation is the coordination among the agents. Basically, there are 3 forms of *coordination strategies*: centralized, decentralized, and multi-level coordination. Centralized coordination assumes a global manager taking care of the coordination. In decentralized coordination, the agents themselves carry this burden. Multi-level coordination, lying in between, can, for instance, assume a manager at every site. Because of the decentralized nature of the Internet, decentralized coordination seems most natural. However, tuning the retrieval performance fancies centralized coordination.

The type of messages sent during communication should carry all the information needed in the protocol used. The Contract Net Protocol, for instance, is not viable since it cannot cope properly with open environments (dynamics of multiple negotiations in ID) and bounded rationality (see [SL95]). An extension by Sandholm and Lesser supplies abstract payment and commitment functions. These will properly have to be filled in for the ID context. Payment can be defined in terms of complexity, efficiency, accuracy, etc. Commitment functions are applied to ensure a certain fairness (for example, to protect users from negotiation violating brokers).

Agents will apply *negotiation strategies* to implement their goal-directed behaviour. The strategies can be based on mathematics, or on heuristics, for instance.

The negotiations can be held according to several *negotiation mechanisms*. Examples are voting (majority, veto, etc.), auctions (English, Dutch, Vickrey, etc.), and contract negotiation. It has to be investigated which mechanisms suit ID: efficiency and privacy are required.

There are *properties of agents* which influence the negotiation. It must be investigated which properties are possible or needed for ID. Examples are benevolence (agents always do what is asked for by absence of conflicting goals), veracity (to communicate false information deliberately), rationality (act to achieve goals), and mobility (agents do not have a stable address).

4 Conclusions

We proposed an economic market paradigm for agent-based ID. It was shown about what issues ID agents can negotiate. The simple IR and IF tasks were set in this framework. Additional aspects of negotiation were touched upon from an ID perspective.

Additional research is needed for further investigation, but also for implementing and testing the ideas. On a theoretical basis, a *classification* of ID system can be made according to the types of negotiation they support. In this way, an ID system can said to be embedded (see [Hui96]) in another if the second system supports at least all the negotiation forms of the first.

References

[BC92] N.J. Belkin and W.B. Croft. Information filtering and information retrieval: Two sides of the same coin? *Communications of the ACM*, 35(12):29–38, December 1992.

[Bru93] P.D. Bruza. *Stratified Information Disclosure: A Synthesis between Information Retrieval and Hypermedia.* PhD thesis, University of Nijmegen, Nijmegen, The Netherlands, 1993.

[Hui96] T.W.C. Huibers. *An Axiomatic Theory of Information Retrieval.* PhD thesis, Department of Computer Science, Utrecht University, November 1996.

[Mar97] S. Marsh. A Community of Autonomous Agents for the Search and Distribution of Information in Networks. In J. Furner and D.J. Harper, editors, *Proceedings of the 19th BCS-IRSG Colloquium on IR Research*, Aberdeen, Scotland, April 1997.

[SL95] T.W. Sandholm and V.R. Lesser. Automated contracting among self-interested bounded rational agents. Technical report, Computing Science Department, University of Massachusetts, Amherst, 1995.

[WBHW97a] B.C.M. Wondergem, P. van Bommel, T.W.C. Huibers, and Th. van der Weide. Towards an Agent-Based Retrieval Engine. In J. Furner and D.J. Harper, editors, *Proceedings of the 19th BCS-IRSG Colloquium on IR research*, Aberdeen, Scotland, April 1997.

[WBHW97b] B.C.M. Wondergem, P. van Bommel, T.W.C. Huibers, and Th.P. van der Weide. Agents in Cyberspace – Towards a Framework for Multi-Agent Systems in Information Discovery. CSI-R9715, Computing Science Institute, University of Nijmegen, Utrecht, The Netherlands, 1997. Submitted to the ERDIS'97 Workshop.

[WJ95] M. Wooldridge and N.R. Jennings. Intelligent Agents: Theory and Practice. *Knowledge Engineering Review*, 10(2):115–152, 1995.

Computing Science Reports

Department of Mathematics and Computing Science
Eindhoven University of Technology

In this series appeared:

96/01	M. Voorhoeve and T. Basten	Process Algebra with Autonomous Actions, p. 12.
96/02	P. de Bra and A. Aerts	Multi-User Publishing in the Web: DreSS, A Document Repository Service Station, p. 12
96/03	W.M.P. van der Aalst	Parallel Computation of Reachable Dead States in a Free-choice Petri Net, p. 26.
96/04	S. Mauw	Example specifications in phi-SDL.
96/05	T. Basten and W.M.P. v.d. Aalst	A Process-Algebraic Approach to Life-Cycle Inheritance Inheritance = Encapsulation + Abstraction, p. 15.
96/06	W.M.P. van der Aalst and T. Basten	Life-Cycle Inheritance A Petri-Net-Based Approach, p. 18.
96/07	M. Voorhoeve	Structural Petri Net Equivalence, p. 16.
96/08	A.T.M. Aerts, P.M.E. De Bra, J.T. de Munk	OODB Support for WWW Applications: Disclosing the internal structure of Hyperdocuments, p. 14.
96/09	F. Dignum, H. Weigand, E. Verharen	A Formal Specification of Deadlines using Dynamic Deontic Logic, p. 18.
96/10	R. Bloo, H. Geuvers	Explicit Substitution: on the Edge of Strong Normalisation, p. 13.
96/11	T. Laan	AUTOMATH and Pure Type Systems, p. 30.
96/12	F. Kamareddine and T. Laan	A Correspondence between Nuprl and the Ramified Theory of Types, p. 12.
96/13	T. Borghuis	Priorean Tense Logics in Modal Pure Type Systems, p. 61
96/14	S.H.J. Bos and M.A. Reniers	The I^2C-bus in Discrete-Time Process Algebra, p. 25.
96/15	M.A. Reniers and J.J. Vereijken	Completeness in Discrete-Time Process Algebra, p. 139.
96/17	E. Boiten and P. Hoogendijk	Nested collections and polytypism, p. 11.
96/18	P.D.V. van der Stok	Real-Time Distributed Concurrency Control Algorithms with mixed time constraints, p. 71.
96/19	M.A. Reniers	Static Semantics of Message Sequence Charts, p. 71
96/20	L. Feijs	Algebraic Specification and Simulation of Lazy Functional Programs in a concurrent Environment, p. 27.
96/21	L. Bijlsma and R. Nederpelt	Predicate calculus: concepts and misconceptions, p. 26.
96/22	M.C.A. van de Graaf and G.J. Houben	Designing Effective Workflow Management Processes, p. 22.
96/23	W.M.P. van der Aalst	Structural Characterizations of sound workflow nets, p. 22.
96/24	M. Voorhoeve and W. van der Aalst	Conservative Adaption of Workflow, p.22
96/25	M. Vaccari and R.C. Backhouse	Deriving a systolic regular language recognizer, p. 28
97/01	B. Knaack and R. Gerth	A Discretisation Method for Asynchronous Timed Systems.
97/02	J. Hooman and O. v. Roosmalen	A Programming-Language Extension for Distributed Real-Time Systems, p. 50.
97/03	J. Blanco and A. v. Deursen	Basic Conditional Process Algebra, p. 20.
97/04	J.C.M. Baeten and J.A. Bergstra	Discrete Time Process Algebra: Absolute Time, Relative Time and Parametric Time, p. 26.
97/05	J.C.M. Baeten and J.J. Vereijken	Discrete-Time Process Algebra with Empty Process, p. 51.
97/06	M. Franssen	Tools for the Construction of Correct Programs: an Overview, p. 33.
97/07	J.C.M. Baeten and J.A. Bergstra	Bounded Stacks, Bags and Queues, p. 15.

61

97/08	P. Hoogendijk and R.C. Backhouse	When do datatypes commute? p. 35.
97/09	Proceedings of the Second International Workshop on Communication Modeling, Veldhoven, The Netherlands, 9-10 June, 1997.	Communication Modeling- The Language/Action Perspective, p. 147.
97/10	P.C.N. v. Gorp, E.J. Luit, D.K. Hammer E.H.L. Aarts	Distributed real-time systems: a survey of applications and a general design model, p. 31.
97/11	A. Engels, S. Mauw and M.A. Reniers	A Hierarchy of Communication Models for Message Sequence Charts, p. 30.
97/12	D. Hauschildt, E. Verbeek and W. van der Aalst	WOFLAN: A Petri-net-based Workflow Analyzer, p. 30.
97/13	W.M.P. van der Aalst	Exploring the Process Dimension of Workflow Management, p. 56.
97/14	J.F. Groote, F. Monin and J. Springintveld	A computer checked algebraic verification of a distributed summation algorithm, p. 28
97/15	M. Franssen	λP-: A Pure Type System for First Order Loginc with Automated Theorem Proving, p.35.
97/16	W.M.P. van der Aalst	On the verification of Inter-organizational workflows, p. 23
97/17	M. Vaccari and R.C. Backhouse	Calculating a Round-Robin Scheduler, p. 23.

02