UNIVERSIDADE DE SÃO PAULO

UMA FERRAMENTA ORIENTADA A LINKS PARA AVALIAÇÃO DE HIPERDOCUMENTOS

RENATA PONTIN DE MATTOS FORTES

Nº 24

NOTAS

Instituto de Ciências Matemáticas de São Carlos

Instituto de Ciências Matemáticas de São Carlos

ISSN - 0103-2577

UMA FERRAMENTA ORIENTADA A LINKS PARA AVALIAÇÃO DE HIPERDOCUMENTOS

RENATA PONTIN DE MATTOS FORTES

Nº 24

NOTAS DO ICMSC
Série Computação

São Carlos
1996

Uma Ferramenta orientada a *links* para Avaliação de Hiperdocumento

Renata Pontin de Mattos Fortes

renata@icmsc.sc.usp.br

Departamento de Ciências de Computação e Estatística

Instituto de Ciências Matemáticas de São Carlos

Universidade de São Paulo

Resumo

Este trabalho apresenta uma ferramenta para exame da estrutura de hiperdocumentos. Essa ferramenta foi aplicada na avaliação do servidor *WWW* do ICMSC-USP, com o objetivo de subsidiar os autores desse hiperdocumento. Os resultados obtidos com a avaliação serão utilizados em um projeto mais abrangente que visa investigar e propor meios de auxílio à autoria para melhorar a qualidade de hiperdocumentos.

Palavras-Chave: Hiperdocumento, Avaliação de Estrutura de Hiperdocumento, Técnicas de Representação de Hiperdocumento e Statecharts.

Abstract

This paper presents a framework to peruse the hyperdocument structure. This framework has been used to evaluate the ICMSC WWW server, which aims to help hyperdocument authors. This evaluation framework is part of a wider project, which intends to investigate and to propose ways to support the authoring process for improving hyperdocument quality.

Keywords: Hyperdocument, Evaluation of Hyperdocument Structure, Techniques for Representing Hyperdocuments and Statecharts.

Índice

1. Introdução **3**

2. Avaliação de Hiperdocumentos **4**

3. A ferramenta orientada a *links* **5**

3.1. Representação de Hiperdocumento da *web* **6**

3.2. A Base de Dados de *links* **10**

3.3. Apresentação dos *links* **10**

4. Aplicação da ferramenta e Resultados obtidos **12**

5. Próximas Etapas **16**

Bibliografia **16**

1. Introdução

Atualmente, a construção de hipertextos tem ocorrido rapidamente nas mais diversas áreas de aplicação. Desde a comunidade científica com sistemas de documentação para ensino e para desenvolvimento de software, até nos ambientes comerciais com sistemas de *marketing* e propaganda, está se configurando como uma nova tecnologia importante aos sistemas de acesso à informações.

A flexibilidade e facilidade de uso dos sistemas hipertexto têm garantido um futuro cada vez mais promissor para sua utilização. Porém, a construção de hiperdocumentos envolvendo a montagem de milhares de nós e de dezenas de milhares de *links* é uma atividade com grande potencial de erros. Além disso, quando a escala do volume de informações cresce bastante, tem sido observado que o leitor pode ficar perdido durante a navegação pelo hiperdocumento.

Muitos trabalhos têm feito propostas para solucionar esse problema, adicionando características aos sistemas hipertexto com diferentes recursos de navegação, tais como: um mapa gráfico (*browser*) para navegação através da estrutura de ligações contida no hiperdocumento e módulos para retroceder *(backtrack)*, incluir marcadores *(bookmarks)*, registrar a história do caminho percorrido pelo leitor. Existem ainda outros trabalhos que propõem soluções para esse problema através de modelagem do hiperdocumento, fornecendo um esquema conceitual apropriado para os nós e *links* do hiperdocumento [Garz-93, Isak-95]. Essa abordagem de modelagem é mais adequada para hiperdocumentos cuja orientação se refira a um determinado domínio de informações, porém como se vê na *WWW* (*World Wide Web*) as páginas e ligações possuem uma dinâmica muito grande de manutenções, o que deve refletir em freqüentes revisões da modelagem para acompanhar essa dinâmica.

De acordo com Brown [Brown-90], o sucesso de um hipertexto está diretamente relacionado à qualidade da autoria. Ele compara o estágio de desenvolvimento de ferramentas para autoria ao estágio da programas dos anos 50.

Neste relatório é apresentada uma ferramenta para exame da estrutura de hiperdocumentos, com a qual foi realizada uma avaliação do servidor da *web*[1] do ICMSC-USP, com o objetivo de subsidiar seus autores (*webmasters* e usuários da Internet que se interessam em colaborar na autoria das páginas da *web*). Essa ferramenta

[1] Para simplificar a leitura será utilizado o termo *web* em lugar de *WWW*.

de avaliação faz parte do trabalho de tese desta autora, cujo objetivo é propor meios de auxílio à autoria visando a melhorar a qualidade de hiperdocumentos.

2. Avaliação de Hiperdocumentos

De acordo com Hatzimanikatis, Tsalidis e Christodoulakis [Hatz-95], não existe uma definição completa de qualidade de hiperdocumento e os métodos e ferramentas que poderiam ajudar na análise e avaliação da qualidade de hiperdocumentos ainda estão em um estágio bastante primitivo com relação ao auxílio efetivo que deveriam fornecer.

A possibilidade de medir os atributos de um hiperdocumento é imprescindível para as atividades de autoria e Engenharia de hipertexto[2]. Os paradigmas apropriados para tal tarefa podem vir da experiência prática nas áreas de Engenharia de Software e de Métricas de Software. Entretanto, a experimentação é uma questão chave para a aplicação de métricas e, para suportá-las, são necessárias ferramentas flexíveis para a obtenção automática das medidas [Hatz-95].

Para a definição de quais medidas são importantes na avaliação de um hiperdocumento há a necessidade de se estabelecer o escopo de suas características. Seguindo o modelo de referência Dexter proposto por Halasz e Schwartz [Hala-90] para sistemas hipertexto, os hiperdocumentos podem ser examinados sob dois ângulos de granularidade distintos: seu conteúdo e estrutural. O modelo Dexter também define a camada de apresentação de sistemas hipertextos, que se refere às formas resultantes de exibição do hiperdocumento em tempo-de-execução, mas essa camada está fora do escopo de avaliação proposto neste trabalho, cuja abordagem é baseada em *links*.

O nível de conteúdo, que é chamado no modelo Dexter de camada interna-dos-componentes (*within component layer*), engloba o conteúdo dos nós e os intra-relacionamentos existentes dentro de um nó, como por exemplo: os parágrafos e os títulos. No nível estrutural, chamado de camada de armazenamento (*storage layer*) no modelo Dexter, está definida a base de dados consistindo da rede de nós e *links*, ou seja, os inter-relacionamentos entre nós e *links*.

Entre os trabalhos relevantes na área de avaliação do conteúdo de documentos encontra-se o de Salton, Allan e Buckley [Salt-94] que visa a definição de *links* automaticamente. Em relação às pesquisas relacionadas à avaliação da estruturação de

---

[2] Termo utilizado por Glushko [Glushko-88]: "necessidade por uma abordagem disciplinada de **Engenharia de Hipertexto**, que trabalhe com a análise da tarefa, análise do documento e considerações cuidadosas das lições obtidas a partir de escritos técnicos e disciplinas de recuperação de informações.

hiperdocumentos, Botafogo e Rivlin et alli [Bota-92, Rivl-94] analisaram a conectividade do grafo que descreve o hiperdocumento e propuseram algoritmos para identificar possíveis agregados, hierarquias e métricas.

A ferramenta proposta e desenvolvida neste trabalho, e que será descrita na próxima seção, parte da identificação dos reusos de componentes de *links*, com base na estruturação do hiperdocumento, com o objetivo de auxiliar os autores a avaliar a organização, estilo e uniformidade dos termos criados para ancoragem dos *links*. Dessa forma, a avaliação suportada pela ferramenta desenvolvida se enquadra entre os dois níveis de granularidade, de conteúdo e de estrutura, uma vez que a ancoragem dos *links* se situa como uma interface entre esses dois níveis, segundo o modelo Dexter.

3. A ferramenta orientada a *links*

A ferramenta orientada a *links* se destina primeiramente aos autores de hiperdocumentos que têm trabalhado com toda a liberdade encontrada em hipertextos como os da *web*, mas, apesar disso, necessitam também de auxílio para gerenciar o volume crescente de informações incorporadas a cada dia. Além disso, é um auxílio para os autores obterem informações mais precisas sobre a organização e o estilo que colocam nos relacionamentos entre as páginas. A ferramenta fornece aos autores uma visão orientada a *links* que os auxiliará a examinar a organização, estilo e uniformidade dos termos utilizados no hiperdocumento.

A ferramenta implementa as seguintes classes de objetos: *link*, nó-fonte, âncora e nó-destino. A Figura 1 esquematiza o modelo de objetos construído.

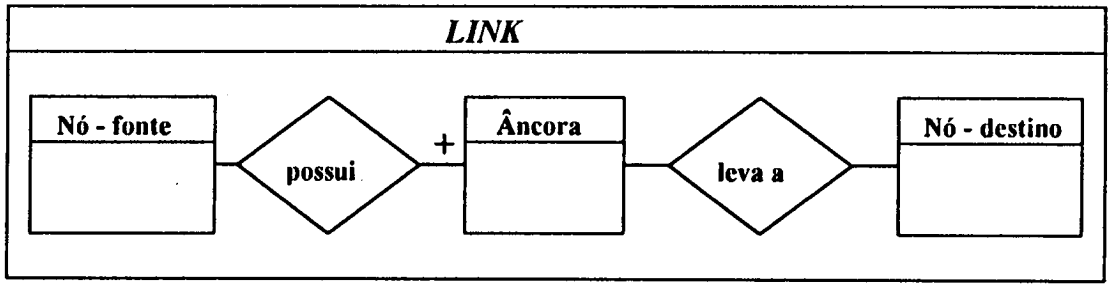

Figura 1 - Modelo de objetos: os objetos componentes do objeto *link*.

Com base nesse modelo de objetos, os módulos funcionais da ferramenta foram definidos, conforme mostra o esquema na Figura 2. Nessa arquitetura foram especificados os seguintes processamentos: representação de sua estrutura de *links*, criação de uma base contendo essas estruturas e apresentação da mesma. Nas próximas seções estão detalhadas as funções de cada um desses processos.

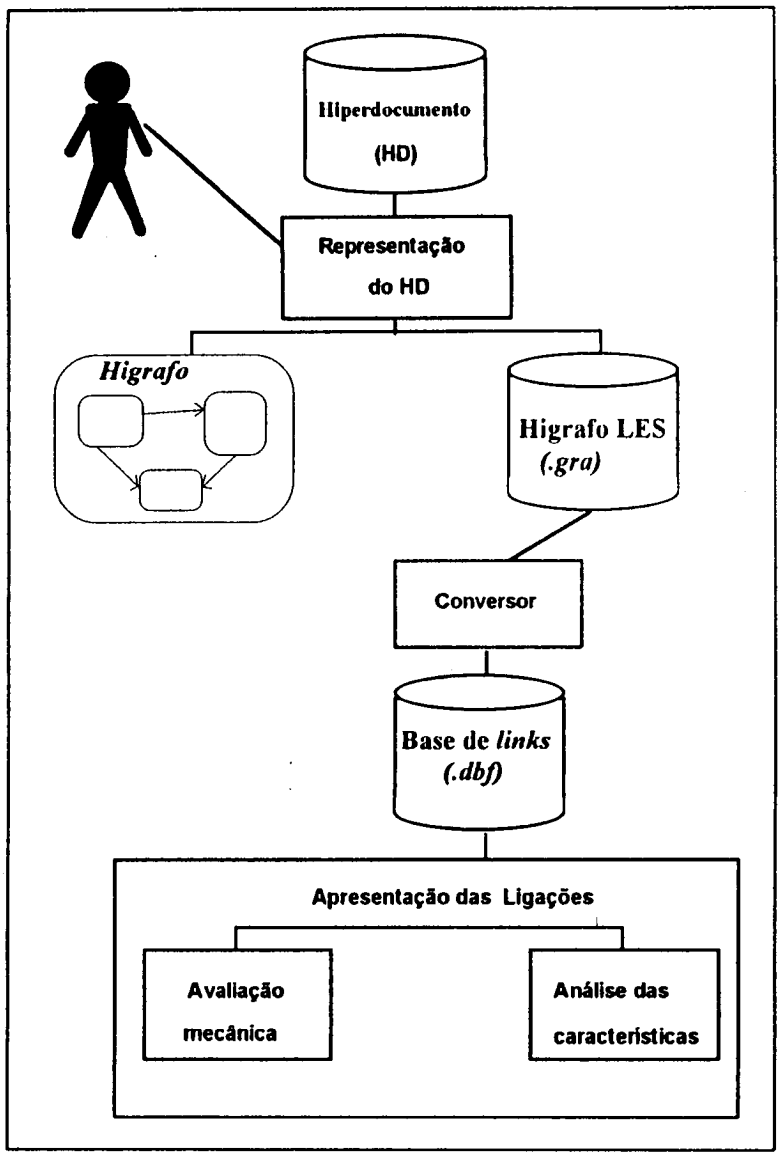

Figura 2 - Módulos de software que compõem a ferramenta orientada a *links*.

3.1. Representação de Hiperdocumento da *web*

O ponto de partida para a construção da ferramenta foi a representação dos *links* do hiperdocumento referente ao sevidor da *web* do ICMSC. Para isso, uma forma natural seria representá-lo através de um grafo dirigido. Mas visando a investigação de abstrações na estrutura do hiperdocumento foi utilizada a técnica Statechart, que é um formalismo visual baseado em *higraph* (*hierarquical graph* ou higrafo em Português), proposto por Harel [Harel-87], e que estende os Diagramas de Transição de Estados (DTE) cuja representação é um grafo dirigido rotulado.

A representação obtida, embora pudesse ser diagramada, era o procedimento inicial de base da ferramenta. Portanto, foi decidido que a representação seria apenas descrita textualmente e foi utilizada a LES (Linguagem de Especificação de Statechart) [Fort-91], que permite a descrição textual dos componentes estruturais de um hiperdocumento, seguindo a sintaxe e semântica dos Statecharts.

```
<html><head><title>ICMSC-USP Homepage</title></head>

<P>     <a href="ICMSC/"><IMG SRC="/fotos/button.gif" ></a>
<a href="ICMSC/">About the Institute</a><br>
Location, organization, departments, courses, research
<p>     <a href="/staff/"><IMG SRC="/fotos/button.gif" ></a>
<a href="/staff/">Faculty </a><br>
List by Departments; links to homepages
<p>     <a href="/ICMSC/icmsc-jr.html"><IMG SRC="/fotos/button.gif" ></a>
<a href="/ICMSC/icmsc-jr.html">ICMSCjr</a><br>
Junior Enterprise of the ICMSC students
<p>     <a href="/eventos/"><IMG SRC="/fotos/button.gif" ></a>
<a href="/eventos/">Events at the Institute</a><br>
Seminars, speeches, symposiums, Summer Courses
<p>     <a href="servicos.html"><IMG SRC="/fotos/button.gif" ></a>
<a href="servicos.html">Services</a><br>
Phone book; WWW, FTP and Gopher servers; news
<p>     <img src="/icones/work.gif" align=left alt="In construction">
<h5 align=center>
<i>Last edited November 06, 1995 <br>
<a href = "mailto:webmaster@icmsc.sc.usp.br">
webmaster@icmsc.sc.usp.br</a></i><p>
<a href=/Portugues/>texto em Portugues </A>
 <a href="http://www.usp.br/">USP Homepage</a>
</h5>
</body>
</html>
```

Figura 3 - A página de entrada do servidor da *web* do ICMSC-USP em HTML.

ICMSC-USP Homepage

☺ About the Institute
Location, organization, departments, courses, research
☺ Faculty
List by Departments; links to homepages
☺ ICMSCjr
Junior Enterprise of the ICMSC students
☺ Events at the Institute
Seminars, speeches, symposiums, Summer Courses
☺ Services
Phone book; WWW, FTP and Gopher servers; news

✍ *Last edited November 06, 1995*
webmaster@icmsc.sc.usp.br

texto em Portugues
USP Homepage

Figura 4 - Apresentação da página de entrada do servidor da *web* do ICMSC-USP, através de um sistema de navegação (*browser*), como o Nescape ou Mosaic.

Como o hiperdocumento da *web* não possui abstrações de composição ou de contexto, o Statechart ficou como um DTE acrescido de um superestado "hiperdoc" abstrato contendo todos os estados (nós) existentes realmente na *web*.

A título de exemplo, na Figura 3 está transcrita parte da página de entrada do ICMSC (*Homepage*), que corresponde ao arquivo "index.html" do servidor da *web* do ICMSC, e na Figura 4 encontra-se esquematizada a apresentação da mesma página, através de um sistema hipertexto. A Figura 5 contém o resultado da execução do módulo, que corresponde à descrição em LES da mesma página exemplificada nas figuras 3 e 4, e na Figura 6 está o seu diagrama Statechart.

O módulo desenvolvido varre o hiperdocumento, seguindo os seus *links* locais, para gerar um arquivo ASCII, com a estrutura do hiperdocumento, ou seja, os seus *links* em LES. O arquivo foi chamado "*dte.gra*". Essa varredura não avança na obtenção dos *links* para arquivos que tivessem que ser obtidos através de algum protocolo de transferência, ficando, portanto, construída uma **representação da estrutura do hiperdocumento local**, ou seja, pertencente ao sistema de arquivos sobre o qual os autores locais tiveram autoria direta.

Para a descrição dos *links* em LES, obtidos dos documentos da *web* em HTML (*Hypertext Markup Language*), foi utilizada a noção de que a âncora que é apresentada ao leitor no nó corrente (nó-fonte) corresponde ao evento que o leitor aciona para que seja seguido o *link*. Portanto, a cadeia de caracteres que fica delimitada pela marcação (*tag*) foi adotada como sendo a âncora-fonte, seja ela a própria a ser apresentada durante a navegação, seja uma composição de um arquivo que possui um *icon* (imagem) que represente a âncora. Por se tratar de *links* unidirecionais, foi feita a correspondência dos nós-fontes como as origens das transições e dos nós-destinos como sendo os estados-destinos das transições.

```
estado "hiperdoc" ou subestados                    "./index.html" def,
"./ICMSC/index.html",                              "../staff/index.html", "..
/ICMSC/icmsc-jr.html",                             "../eventos/index.html",
"./servicos.html",                                 "mailto:webmaster@icmsc.sc.usp.br",
"../Portugues/index.html",                         "http://www.usp.br/"

estado "./index.html" atomo                        estado "./ICMSC/index.html" atomo
estado "../staff/index.html" atomo                 estado "../ICMSC/icmsc-jr.html" atomo
estado "../eventos/index.html" atomo               estado "./servicos.html" atomo
estado "mailto:webmaster@icmsc.sc.usp.br" atomo    estado "../Portugues/index.html" atomo
estado "http://www.usp.br/" atomo  ;

evento "<IMG SRC='/fotos/button.gif'>"             origem "./index.html"
   destino "./ICMSC/index.html"
evento "About the Institute"                       origem "./index.html"
   destino "./ICMSC/index.html"
evento "<IMG SRC='/fotos/button.gif'>"             origem "./index.html"
   destino "../staff/index.html"
evento "Faculty"                                   origem "./index.html"
```

```
   destino "../staff/index.html"
evento "<IMG SRC='/fotos/button.gif'>"          origem "./index.html"
   destino "../ICMSC/icmsc-jr.html"
evento "ICMSCjr"                                origem "./index.html"
   destino "../ICMSC/icmsc-jr.html"
evento "<IMG SRC='/fotos/button.gif'>"          origem "./index.html"
   destino "../eventos/index.html"
evento "Events at the Institute"                origem "./index.html"
   destino "../eventos/index.html"
evento "<IMG SRC='/fotos/button.gif'>"          origem "./index.html"
   destino "./servicos.html"
evento "Services"                               origem "./index.html"
   destino "./servicos.html"
evento "webmaster@icmsc.sc.usp.br"              origem "./index.html"
   destino "mailto:webmaster@icmsc.sc.usp.br"
evento "texto em Portugues"                     origem "./index.html"
   destino "../Portugues/index.html"
evento "USP Homepage"                           origem "./index.html"
   destino "http://www.usp.br/" ;
```

Figura 5 - Declarações em LES da estrutura do hiperdocumento composto da *Homepage* (página de entrada) e seus *links*, do servidor da *web* do ICMSC-USP.

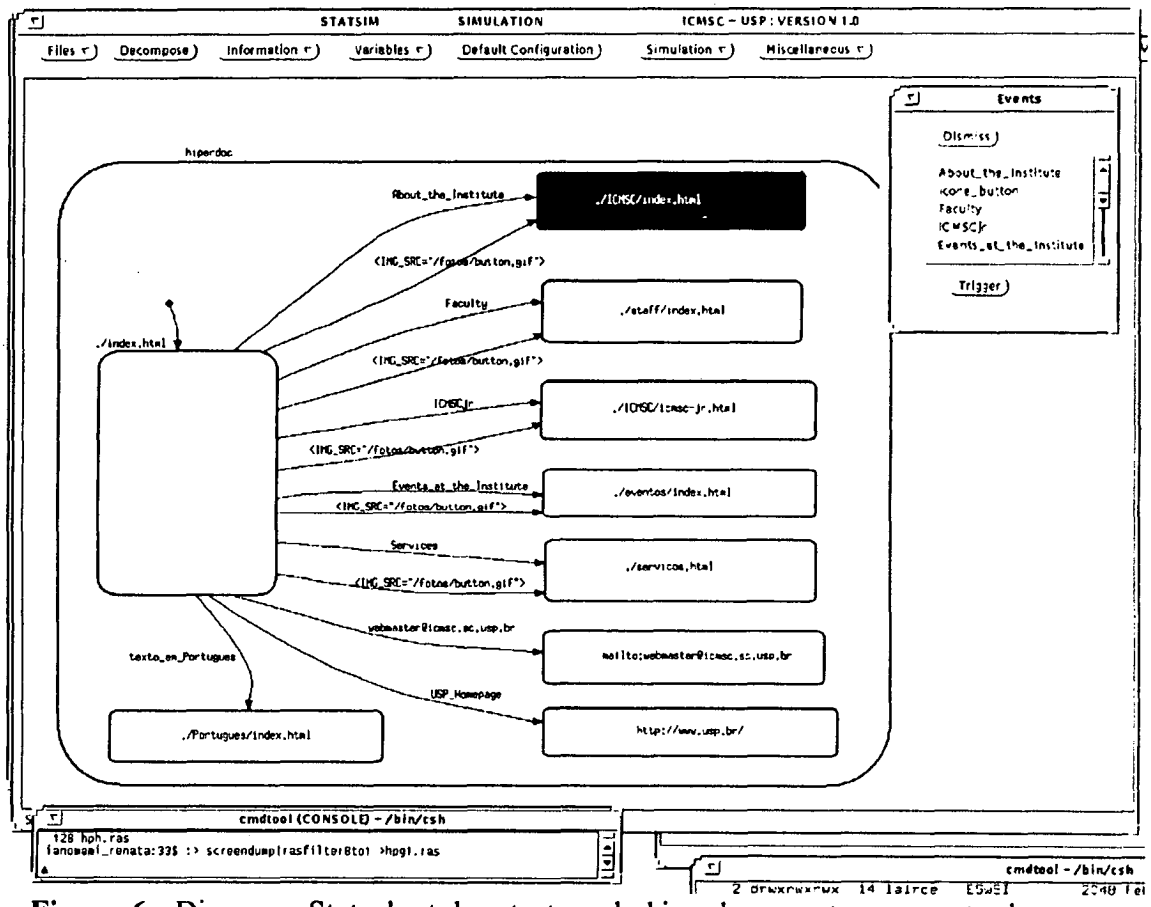

Figura 6 - Diagrama Statechart da estrutura do hiperdocumento composto da *Homepage* e seus *links*, do servidor da *web* do ICMSC-USP que foi obtido de um *dump* de tela produzida no ambiente StatSim (*Statechart Simulator*).

3.2. A Base de Dados de *links*

Obtidos os elementos que compõem a estrutura do hiperdocumento, no processamento de "Representação do HD", na forma de um arquivo ASCII, cada *link* ficou mapeado na forma de componentes: evento-origem-destino descritos em LES, que correspondem às transições do Statechart.

O módulo "Conversor" é responsável pela criação de identificadores únicos para os componentes de *links*, em adequação ao modelo de objetos representado na figura 1 e para suprir as consultas disponibilizadas no módulo de "Apresentação dos *links*". No caso, foi utilizado o software Microsoft Visual FoxPro versão 3.0 (um ambiente de desenvolvimento de programas com suporte de um Sistema Gerenciador de Base de Dados Relacional) que já possui um conversor automático interno de bases de dados em ASCII e gera a tabela correpondente em formato de base de dados relacionais (*.dbf*).

Por se tratar da estruturação de páginas escritas em HTML, as tuplas referentes aos *links* puderam ser assim definidas:

(id-link, âncora-fonte, nó-fonte, nó-destino.âncora-destino)

onde : id-link = identificador-de-link (chave- primária);
âncora-fonte = evento;
nó-fonte = origem;
nó-destino.âncora-destino = destino.

A âncora-destino foi considerada como pertencente ao atributo do nó-destino porque ela simplesmente desempenha um papel de desvio dentro do ponto de apresentação do nó-destino quando um usuário-leitor decide seguir o *link* que a contenha. Dessa forma, todos os *links* que continham mesmos nós-destinos, mas com âncoras-destino distintas, foram considerados *links* distintos. Além disso, por HTML suportar somente a definição de *links* unidirecionais, a âncora-destino não possui o papel de controle, como acontece com a âncora-fonte, que é o ponto de entrada para a interação durante a navegação de um usuário leitor.

3.3. Apresentação dos *links*

A proposta da ferramenta desenvolvida é fornecer ao autor uma visão complementar de sua autoria, enfatizando a apresentação dos *links* criados no hiperdocumento. Ao invés do autor ver somente os *links* a partir dos nós que ele cria, a abordagem está em focalizar os *links* e então os nós que eles conectam.

Nessa visão orientada a *links* é apresentada uma relação geral de todos os *links* que o hiperdocumento possui, para serem examinados. O autor obtém a visão de todos

os componentes estruturais do hiperdocumento na forma de uma lista, uma vez que a montagem da estrutura toda em mapas gráficos ou diagramáticos discutida na literatura em geral é uma tarefa complexa e geralmente leva a um enorme emaranhado de linhas de conexão.

Mas o ponto importante é que para esse exame a ser feito pelo autor, a ferramenta funciona como um hipertexto dedicado, onde a obtenção dos nós, fonte e destino que participam de um *link*, pode ser feita ao selecioná-lo na lista apresentada.

A janela principal possui uma lista de todos os nomes de âncoras que o autor usou no hiperdocumento. É importante lembrar que no modelo de referência de Dexter [Hala-90], as âncoras possuem um tratamento específico e possuem para tanto uma camada de interface entre a camada de armazenamento e interna aos componentes, chamada de "Ancoragem". Vale ressaltar, portanto, que a Ancoragem é um mecanismo diferencial dos sistema hipertexto e que também deve ser avaliada para atender e contribuir com a qualidade de um hiperdocumento.

Essa relação contendo os nomes de âncoras fica disponível para seleção, e uma vez selecionado um *link* através de um determinado nome de âncora, são obtidas duas novas listas (em janelas laterais à janela principal): a primeira, com todos os nós-fontes que possuem a âncora selecionada como ponto de partida para o *link* e a segunda, com todos os nós-destinos que possam ser alcançados como ponto de chegada dos *links* que partiram da âncora selecionada com aquele nome apresentado como indicador de presença de *link*.

Dessa forma, o autor pode acionar um *link* e visualizar todos os componentes estruturais envolvidos com aquela apresentação de âncora como ponto de interação.

A navegação convencional em um sistema hipertexto é efetivada através das âncoras dos nós-fontes[3], uma vez que elas são os elementos de controle através dos quais os *links* ficarão disponíveis ou habilitados para interação, durante a navegação. Como o *link* é um tipo de relacionamento **n** para **m** (ou seja, do ponto de vista do identificador de *link*, a mesma âncora-fonte pode rotular **n** nós-fontes a **m** nós-destinos), com este módulo de "Apresentação de *links*" o autor poderá ver as diferentes conexões que as âncoras-fontes estão rotulando no hiperdocumento. O resultado dessa apresentação permite ao autor analisar as inconsistências geradas por nomeação indevida dos *links*.

---

[3] Para simplificar a leitura, doravante serão chamadas de **âncoras-fonte** os nomes das âncoras dos nós-fontes.

Na Figura 7 encontra-se um *dump* de tela com as janelas que fazem a apresentação dos *links*. Este módulo de apresentação foi prototipado utilizando-se a linguagem TCL na versão 7.4 com o *toolkit* gráfico TK versão 4.0. Para o armazenamento dos dados foi utilizado um servidor SQL "Mini-SQL" versão 1.0.6 instalado numa máquina remota.

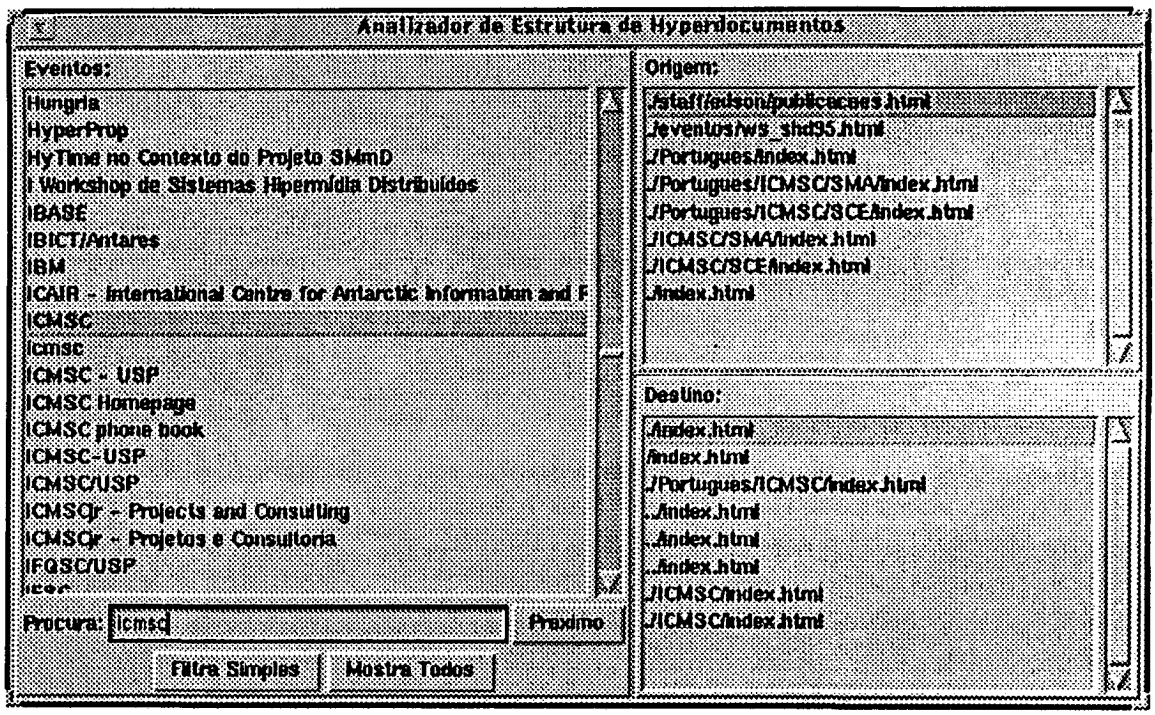

Figura 7 - As janelas que apresentam os *links* de um Hiperdocumento da *web*, mostrando os nós-fonte e destino que são conectados por eles.

4. Aplicação da Ferramenta e Resultados obtidos

A partir dos procedimentos iniciais para representação da estrutura geral do hiperdocumento, pôde-se notar que análises mais específicas para avaliação de hiperdocumentos poderiam ser feitas. Um dos pontos notados foi o da liberdade de nomeação de *links*. A máquina de hipertexto requer endereços precisos para localizar os nós-destino durante a navegação. Os mecanismos para resolução de endereços na busca para recuperação dos nós-destino podem ser:

→ mecanismo antecipado - o texto do nó-fonte está previamente compilado, ficando disponíveis para o sistema de navegação os endereços dos nós-destinos já resolvidos;

→ mecanismo de interpretação - através de um protocolo de comunicação (como é o caso do http que interpreta os arquivos-textos dos nós-fontes em HTML na *web*, os quais deixam os endereços especificados em URL - *Universe Resource Location*), onde os endereços estão completamente especificados.

Portanto, enquanto a nomeação dos nós está consolidada, o mesmo não acontece com a nomeação dos *links*, que fica inteiramente livre e a cargo do autor. É nesse ponto que o autor usa de sua criatividade para imprimir o estilo que pretende para o hiperdocumento.

Outro ponto importante para avaliação foi o de reuso de componentes de *links*. Por exemplo, em um levantamento das relações existentes entre os *links* (nó-fonte/ âncora-fonte/ nó-destino), seria possível ter casos configurados quanto ao reuso de nó-fonte, âncora-fonte ou nó-destino em outros *links*. Esses casos puderam ser esquematizados conforme mostra a Figura 8.

Aqui está-se lidando com um domínio de informações sobre documentos e, portanto, os casos esquematizando o reuso de nós e *links* conduzem a avaliações genéricas, indicando possíveis inconsistências ou informações a serem investigadas. Na Tabela 1 são apresentados os primeiros resultados obtidos com relação ao levantamento de incidência dos casos 0 a 7, esquematizados na Figura 8, no hiperdocumento do servidor da *web* do ICMSC de outubro de 1995.

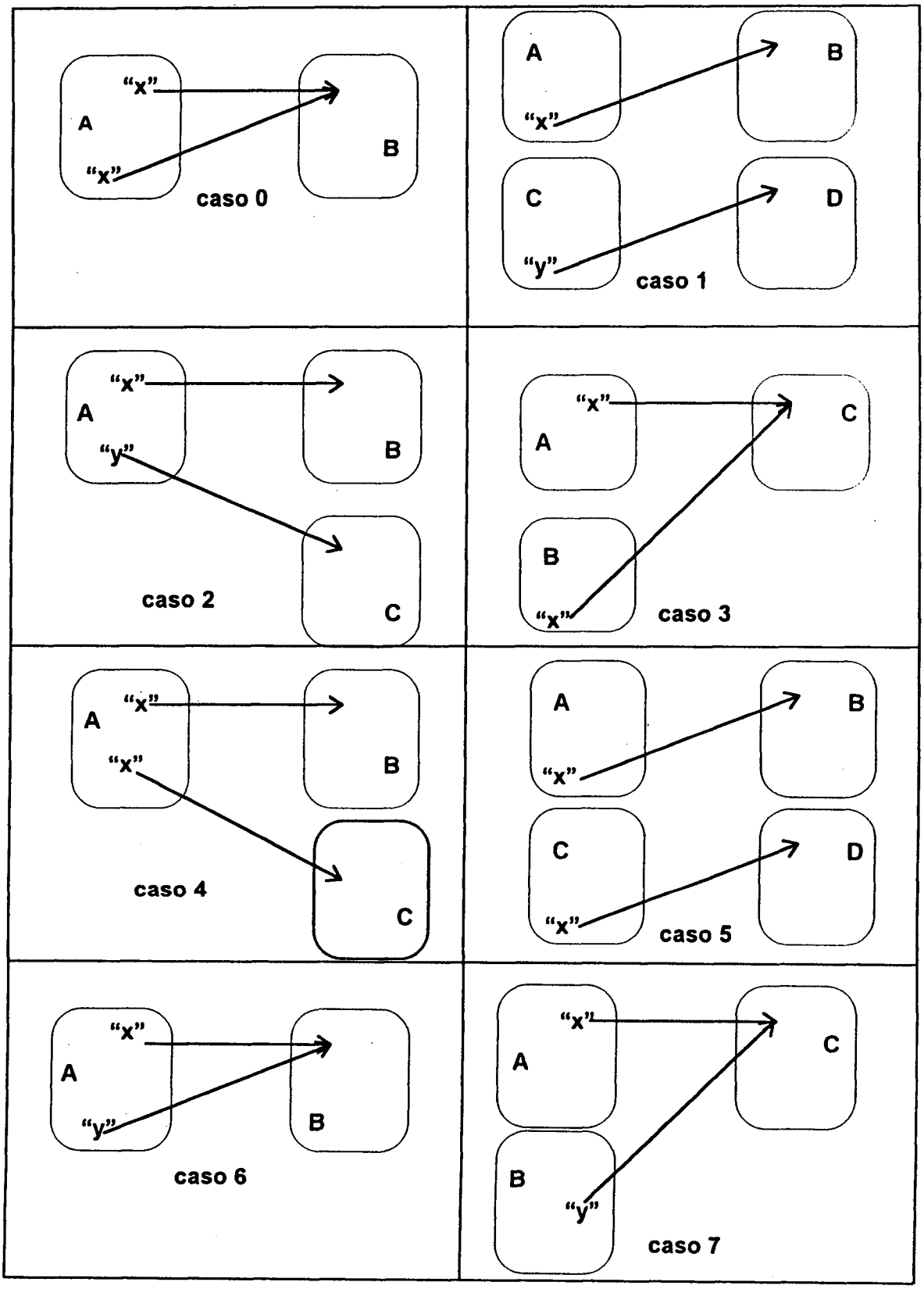

Figura 8 - Casos de *links* considerando-se o reuso de nós / âncoras-fonte.

Tabela 1 - Primeiros resultados das *queries* sobre os casos de *links* na *web*.

caso	questão	resultados (*links*)
0	*Links* que possuem mesma âncora-fonte, mesmo nó-fonte e mesmo nó-destino?	140 5,99%
1	*Links* que não possuem qualquer reuso de estrutura em relação aos demais?	19 0,81%
2	*Links* que partem de um mesmo nó-fonte, com âncoras-fonte distintas e chegam em nós-destino distintos?	1.674 71,72%
3	*Links* que possuem mesma âncora-fonte e chegam no mesmo nó-destino?	649 27,80%
4	*Links* com mesma âncora-fonte que partem do mesmo nó-fonte e chegam em nós-destino distintos?	182 7,79%
5	*Links* que possuem uma mesma âncora-fonte, que partem de nós-fonte diversos e chegam em nós-destino distintos?	281 12,03%
6	Âncoras-fonte distintas que partem de mesmo nó-fonte e chegam a um mesmo nó-destino?	216 9,25%
7	*Links* que chegam a um mesmo nó-destino, a partir de âncoras-fonte distintas?	413 17,69%

A relação "**dte.dbf**" possuia 2.334 registros (obtidos a partir de 1.406 nós), que corresponde ao total de *links* obtidos da *web* local naquela época. Cada relação intermediária resultante de *query* em SQL foi contabilizada para apresentar as porcentagens de incidência de cada caso na *web*.

Deve-se observar que esse levantamento de casos possui degenerações que estão embutidas. Por exemplo, quando o hiperdocumento todo é composto de um nó somente os *links* seriam representados através de setas saindo e chegando naquele nó. Uma consideração feita foi de que as âncoras-destino, quando diferentes, seriam representadas como pertencentes a nós-destino distintos. Portanto, os casos 1, 2, 4 e 5 foram contabilizados considerando também as incidências de âncoras-destino distintas mas contidas no mesmo nó-destino.

Nota-se, na Tabela 1, que os casos de maior incidência (2 e 3) são os que representam páginas índice (com muitos *links* de saída) e páginas bastante referenciadas (com muitos *links* de chegada), respectivamente, são os casos mais típicos de definição de *links* e revelam certa similaridade com escrita impressa..

Já o caso 4, por exemplo, em que ocorre de um nó-fonte com duas âncoras-fonte, de mesmo nome nele contidas, dispararem *links* para nós-destino distintos, indica à primeira vista uma inconsistência. No entanto, sua incidência não é desprezível (~8%). Deve-se analisar então, as suas ocorrências. Considerando-se o "contexto" da âncora-

fonte, o caso 4 pode ser bem empregado, demonstrando uma uniformidade na nomenclatura. Por exemplo, a âncora-fonte "Exemplo", que aplicada a diferentes regiões no texto do nó-fonte direciona a nós-destino distintos contendo exemplos relacionados ao contexto de onde a âncora-fonte possuía seu ponto de partida. Assim, pode indicar que o tipo de *link* com âncoras-fonte iguais leva a uma classe de nós "*exemplo*", e que poderia ser uma abstração de agregações de nós-exemplo.

Com a visão de todos os *links* oferecida pela ferramenta desenvolvida neste trabalho, tem-se portanto, um auxílio para os autores de hiperdocumento *web* que permite examinar sua estrutura e avaliar a organização, estilo e uniformidade dos termos criados para ancoragem dos *links*.

5. Próximas Etapas

Os estudos envolvidos durante o desenvolvimento da ferramenta apresentada aliados aos resultados obtidos pela aplicação da mesma à *web* do ICMSC - USP foram importantes no sentido de possibilitar que um conhecimento maior sobre a autoria de hiperdocumentos fosse adquirido, principalmente em relação ao processo de definição dos *links*..

Os próximos passos desta pesquisa incluem a aplicação da ferramenta em diferentes *sites* de *webs* e a investigação com mais detalhes das características de reuso de *links*, visando a obtenção de abstrações das possíveis agregações em função de relacionamentos entre nós por reuso de componentes de *links*.

Encontra-se em fase final a descrição formal dos casos de reuso de componentes de *links* bem como uma classificação de suas aplicações, com o objetivo de subsidiar os autores de *webs* sobre a correta avaliação desse reuso.

Bibliografia

[Bota-92] BOTAFOGO,R.A; SHNEIDERMAN,B. "Structural Analysis of Hypertexts: Identifying Hierarchies and Useful Metrics." *ACM Transactions on Information Systems*, v.10, n.2, p.142-80, April 1992.

[Brown-90] BROWN,P.J. - Assessing the quality of Hypertext documents - In:RIZS,A.; STRITZ,N. & ANDRÉ,J., eds. *Hypertext: Concepts, Systems and Applications* Cambridge, Great Britain, Cambridge University Press, 1990. /*Proceedings of* ECHT'90, Versailles, France, November 1990./ p.1-12.

[Fort-91] FORTES, R.P.M.; MASIERO, P.C. - *Uma Linguagem para Especificação de Statecharts e seu Analisador*. Documento de Trabalho, Cód. ME-91/01, ICMSC-USP, 1991.

[Garz-93] GARZOTTO,F.; PAOLINI,P.; SCHWABE,D. "HDM - A Model-Based Approach to Hypertext Application Design" *ACM Transactions on Information Systems*, v.11, n.1, p.1-26, January 1993.

[Hala-90] HALASZ,F. ; SCHWARTZ,M. "The Dexter hypertext reference model" In: MOLINE,J.; BENIGNI,D.; BARONAS,J., eds. *Proceedings of the NIST Hypertext Standardization Workshop* Gaithersburg, January 1990. National Institute of Standards and Technology Special Publication 500-178. Washington: U.S. Government Printing Office, 1990. p.95-133.

[Hare-87] HAREL, D. - STATECHARTS: a Visual Formalism for Complex Systems. *Science of Comp. Programming*, 8, p. 231-274, 1987.

[Hatz-95] HATZIMANIKATIS,A.E.; TSALIDIS,C.T.; CHRISTODOULAKIS,D. "Measuring the Readability and Maintainability of Hyperdocuments" *Software Maintanance: Research and Practice*, v.7, p.77-90, 1995.

[Isak-95] ISAKOWITZ,T.; STOHR,E.A., BALASUBRAMANIAN,P. "RMM: A Methodology for Structured Hypermedia Design." *Communications of the ACM*, v.38, n.8, p.34-44, August 1995.

[Rivl-94] RIVLIN,E; BOTAFOGO,R.; SHNEIDERMAN,B. "Navigating in Hyperspace: Designing a Structure-based toolbox" *Communications of the ACM*, v.37, n.2, p.87-96, February 1994.

[Salt-94] SALTON,G.; ALLAN,J.; BUCKLEY,C. "Automatic Structuring and Retrieval of Large Text Files" *Communications of the ACM*, p.97-108, February 1994.

NOTAS DO ICMSC

SÉRIE COMPUTAÇÃO

023/96 BRIGANTE, W.J.; MOREIRA, E.S. - Utilização de Monitores OLTP no gerenciamento de ambientes de manufatura integrada voltados à produção discreta.

022/96 CANSIAN, A.M.; MOREIRA, E.S.; MAURO, R.B.; MORISHITA, F.T.; CARVALHO, A.C.P.L.F. - Um sistema adaptativo de detecção de intrusão em redes de computadores.

021/95 BEZERRA, L.A.F.; SANTANA, R.H.C.; SANTANA, M.J. - Sistema auxiliar de arquivos baseado em disco WORM para ambientar computacional distribuído.

020/95 NUNES, M.G.V.; HASEGAWA, R. - PROTEMA: intelligent tutoring systems for mathematics.

019/95 OLIVEIRA, M.C.; TURINE, M.A.S.; MASIERO, P.C. - A statechart - based model for hypertext.

018/95 PIMENTEL, M.G.C. - Alternative operations for browsing hypertext.

017/94 ROMEIRO, N.M.L.; CASTELO FILHO, A. - Análise Comparativa de Métodos Numéricos de equações algebrico-diferenciais.

016/94 MAGALHÃES, A.L.C.C.; SIQUEIRA, M.F.; OLIVEIRA, M.C.F. - Operadores de Euler na modelagem por fronteira: conceito, aplicação, estudos de casos.

015/94 ODA, C.S.; MOREIRA, E.S. - ASNMP graphical network monitor with automatic topology discovery.

014/94 FELIPE, L.S.G.; FRANCO, N.M.B. - Sobre a ordem de convergência para as equações integrais de volterra de segunda espécie tipo Abel com soluções não suaves.