

Pontifícia Universidade Católica do Rio Grande do Sul
Faculdade de Informática
Programa de Pós-Graduação em Ciência da Computação

**Um Algoritmo Determinístico para Adaptação
Automática de Documentos de Conteúdo Variável**

Ricardo Farias Bidart Piccoli

**Dissertação apresentada como
requisito parcial à obtenção do
grau de mestre em Ciência da
Computação**

Orientador: Prof. Dr. João Batista Souza de Oliveira

Porto Alegre
2010

Dados Internacionais de Catalogação na Publicação (CIP)

P591a Piccoli, Ricardo Farias Bidart
Um algoritmo determinístico para adaptação automática de documentos de conteúdo variável / Ricardo Farias Bidart Piccoli.
– Porto Alegre, 2009.
102 f.

Diss. (Mestrado) – Fac. de Informática, PUCRS
Orientador: Prof. Dr. João Batista Souza de Oliveira

1. Informática. 2. Algoritmos. 3. Documentos Digitais.
4. Estrutura de Dados. I. Oliveira, João Batista Souza de.
II. Título.

CDD 005.73

**Ficha Catalográfica elaborada pelo
Setor de Tratamento da Informação da BC-PUCRS**



Pontifícia Universidade Católica do Rio Grande do Sul
FACULDADE DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

TERMO DE APRESENTAÇÃO DE DISSERTAÇÃO DE MESTRADO

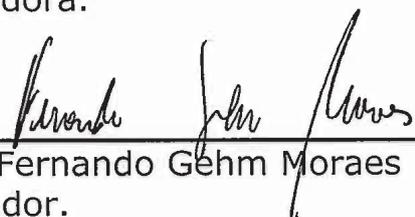
Dissertação intitulada "**Um Algoritmo Determinístico para Adaptação Automática de Documentos de Conteúdo Variável**", apresentada por Ricardo Farias Bidart Piccoli, como parte dos requisitos para obtenção do grau de Mestre em Ciência da Computação, Sistemas Interativos de Visualização, aprovada em 23/01/09 pela Comissão Examinadora:


Prof. Dr. João Batista Souza de Oliveira - PPGCC/PUCRS
Orientador


Prof. Dr. Márcio Sarroglia Pinho - PPGCC/PUCRS


Prof.ª Dra. Luciana Porcher Nedel - UFRGS

Homologada em 30/03/10, conforme Ata No. 005 pela Comissão Coordenadora.


Prof. Dr. Fernando Gehm Moraes
Coordenador.

PUCRS

Campus Central

Av. Ipiranga, 6681 - P32 - sala 507 - CEP: 90619-900
Fone: (51) 3320-3611 - Fax (51) 3320-3621
E-mail: ppgcc@pucrs.br
www.pucrs.br/facin/pos

Ao meu avô Elbio (in memoriam).

*“The art of simplicity is a puzzle of
complexity.”*

Douglas Horton

Agradecimentos

À minha família, principalmente aos meus pais e avós, não somente pelo apoio financeiro e moral, mas por todo amor, compreensão nas horas difíceis e dedicação com a minha criação como ser humano.

Ao meu orientador e professor João Batista, por ter sido – durante estes sete anos de convivência – um modelo a seguir como profissional e ser humano, com toda a paciência, criatividade, humildade e ética que lhe são característicos. Em minha opinião, é uma das pouquíssimas pessoas em nossa área que domina a arte de solucionar problemas de forma inovadora (o famoso “*ahá*” de todo gênio), e por isso me orgulho de ser seu aprendiz.

À minha namorada Andriele – um presente de Deus enviado durante o mestrado –, que por trás de sua fragilidade aparente se revela uma mulher que, com coragem e determinação ímpares, soube superar as tragédias da vida. Ela é e sempre será minha razão de continuar e onde encontrarei força e paz de espírito.

Aos meus grandes amigos e colegas que fiz na universidade:

- Japa, Maurício, Felipe, Aline e a profa. Isabel, pela cooperação no trabalho e companheirismo, mesmo nas horas em que fui difícil de lidar;
- Thais, Ana Karpouzas, Tiago Silva, Machado, Cynthia, Fernando, Leonardo, Karine, entre os mais presentes de muitos dos amigos e colegas que me deram apoio moral e psicológico no mestrado;
- Márcio, Mateus, Thiago, grandes companheiros desde o começo da faculdade, tenho certeza de que estarão entre os melhores profissionais de nossa geração.

Finalmente, à todos que não lembrei ou citei aqui, mas que sabem o papel que tiveram no meu crescimento pessoal e profissional, tenho a honra de tê-los em minha vida.

Resumo

Este trabalho descreve um novo algoritmo para diagramação (i. e., *layout*) automática de documentos de conteúdo variável. O objetivo é mostrar que é possível obter documentos de alta qualidade eficientemente a partir de conteúdos dinâmicos e variáveis (por exemplo, notícias obtidas da *internet*, catálogos de produtos, contas, entre outros), que é uma tendência atual na produção de documentos personalizados. A abordagem aqui descrita, ao contrário da maioria dos trabalhos publicados nesta área, não utiliza métodos aproximativos (por exemplo, algoritmos genéticos) para associar regiões de uma página com os elementos de um documento. A escolha de um algoritmo determinístico foi feita devido às necessidades de desempenho e confiabilidade exigidas no contexto de *Variable Data Printing*, que é uma das aplicações deste trabalho. O método funciona através da realização de divisões sucessivas em uma página, alocando regiões de acordo com a área exigida para cada elemento (conteúdo), ao invés de utilizar geometrias fixas para os mesmos como em outros trabalhos nesta área. A escolha de como dividir a página é baseada em uma heurística simples, capaz de obedecer alguns critérios estéticos importantes em *design* de documentos como alinhamento, regularidade, etc. bem como a ordem de leitura definida pelos elementos do documento e a possibilidade de agrupar elementos relacionados entre si. Além disso, o algoritmo também é capaz de distribuir conteúdo em documentos compostos por várias páginas. Como estudo de caso, o algoritmo foi desenvolvido para produzir diagramações de documentos similares a capas de jornal, tanto no formato quanto no estilo. Finalmente, este trabalho apresenta conceitos relacionados a documentos adaptativos, uma revisão rápida dos principais trabalhos publicados nesta área, o algoritmo desenvolvido em detalhes, assim como uma avaliação empírica dos resultados, considerando aspectos qualitativos dos documentos produzidos e, em menor extensão, o desempenho do algoritmo proposto.

Palavras-chave: Engenharia de Documentos. Documentos de Conteúdo Variável. Diagramação Automática de Documentos. VDP. Algoritmos de Divisão e Conquista. Algoritmos Gulosos.

Abstract

This work describes a new algorithm for automatic layout of variable content documents. The aim is to demonstrate that high-quality page designs can be efficiently obtained from dynamic and variable content (such as content from the internet, sales brochures, bills, and others), which is a current trend in the production of personalized documents. The approach described here, contrary to other works published in this field, does not make use of approximative methods (such as genetic algorithms) to assign regions of a page to the contents of a document. The choice of a deterministic algorithm was made due to the performance and reliability requirements in the context of Variable Data Printing, which is one of the applications of this work. The method works by making successive divisions on a page, allocating regions according to the areas required by each element (content), instead of constraining their geometries as in previous works in this field. The choice of how to divide a page is based on a simple heuristic, capable of preserving some important document design criteria such as alignment, homogeneity, etc. as well as the reading order provided by the input and the possibility of grouping related items. Besides that, the algorithm is also capable of distributing content in a document that is composed by several pages. As a case study, the algorithm was developed to produce document layouts that are similar to newspapers covers, both in design and style. Finally, this work presents concepts regarding adaptive documents, a quick survey of the most relevant works in this field, a detailed description of the algorithm that was developed, as well as an empirical evaluation of the results, regarding qualitative aspects from the generated documents and, to a lesser extent, the performance of the proposed algorithm.

Keywords: Document Engineering. Variable Information Documents. Automatic Document Layout. Variable Data Printing. Divide-and-conquer Algorithms. Greedy Algorithms.

Lista de Figuras

Figura 1	<i>Workflow</i> de um ambiente de VDP.	28
Figura 2	Um exemplo de <i>template</i> de documento de conteúdo variável e quatro instâncias geradas a partir do mesmo [21].	29
Figura 3	Um exemplo de documento cujo <i>layout</i> foi gerado automaticamente a partir do conteúdo, e a readaptação do mesmo documento após a inserção de uma nova imagem como conteúdo.	31
Figura 4	Componentes de uma página de documento utilizado como entrada para o algoritmo, com 6 colunas neste exemplo.	41
Figura 5	Ilustração do <i>workflow</i> utilizado para validação do algoritmo desenvolvido.	43
Figura 6	Exemplo dos arquivos gerados no processo de geração do documento final.	43
Figura 7	<i>Workflow</i> detalhado do algoritmo desenvolvido.	45
Figura 8	Ilustração do algoritmo e a árvore gerada pelo processo de bissecção.	49
Figura 9	Um exemplo do algoritmo de bissecção em uma página de documento.	52
Figura 10	Exemplo do emprego de colunas visuais em um documento com um número fixo de colunas lógicas.	53
Figura 11	Pós-processamento nas regiões de texto para garantir a homogeneidade da página. Neste caso, o tamanho da fonte e espaçamento entre linhas é aumentado até que o texto ocupe toda a região.	54
Figura 12	Exemplo de preservação da ordem de leitura	55
Figura 13	Um documento de uma página contendo 10 elementos.	56
Figura 14	Um documento gerado com 10 elementos dividido em duas páginas.	56
Figura 15	Um documento gerado com 10 elementos dividido em três páginas.	57
Figura 16	Um documento gerado com 10 elementos dividido em quatro páginas.	57
Figura 17	Um documento sem grupos ilustrando a desconexão visual entre o texto e a imagem no canto superior da página.	60
Figura 18	Processo de agrupamento de elementos em uma lista.	60
Figura 19	O mesmo documento da Figura 17, porém com agrupamento lógico entre elementos.	61
Figura 20	Resultados gerados pela avaliação experimental.	67
Figura 21	Gráficos para a avaliação de performance do algoritmo de diagramação automática.	68
Figura 22	Caso de teste número 1.	78
Figura 23	Caso de teste número 2.	79
Figura 24	Caso de teste número 3.	80
Figura 25	Caso de teste número 4.	81
Figura 26	Caso de teste número 5.	82
Figura 27	Caso de teste número 6.	83

Figura 28	Caso de teste número 7.	84
Figura 29	Caso de teste número 8.	85
Figura 30	Caso de teste número 9.	86
Figura 31	Caso de teste número 10.	87
Figura 32	Caso de teste número 11.	88
Figura 33	Caso de teste número 12.	89
Figura 34	Caso de teste número 13.	90
Figura 35	Caso de teste número 14.	91
Figura 36	Caso de teste número 15.	92
Figura 37	Caso de teste número 16.	93
Figura 38	Caso de teste número 17.	94
Figura 39	Caso de teste número 18.	95
Figura 40	Caso de teste número 19.	96
Figura 41	Caso de teste número 20.	97
Figura 42	Formulário para participação dos voluntários na avaliação empírica deste trabalho (Capítulo 5).	100
Figura 43	Questionário pré-teste para avaliação empírica (Capítulo 5).	101
Figura 44	Tabela para avaliação empírica dos casos de teste (Capítulo 5).	102

Lista de Tabelas

Tabela 1	Áreas dos elementos em uma página.	56
Tabela 2	Áreas divididas em duas páginas.	56
Tabela 3	Áreas divididas em três páginas.	57
Tabela 4	Áreas divididas em quatro páginas.	57
Tabela 5	Casos de teste para comparação por medidas estéticas.	64
Tabela 6	Dados da amostra coletados no questionário pré-teste (Figura 43). . . .	66
Tabela 7	Médias e modas das respostas dos voluntários para cada caso de teste, incluindo a respectiva versão reordenada do documento.	66

Lista de Siglas

PDA	<i>Personal Digital Assistant</i>
VDP	<i>Variable Data Printing</i>
RSS	<i>Really Simple Syndication</i>
VID	<i>Variable Information Documents</i>
PPML	<i>Personalized Print Markup Language</i>
XML	<i>eXtensible Markup Language</i>
PDF	<i>Portable Document Format</i>

Sumário

1	Introdução	23
1.1	Motivação	23
1.1.1	Documentos dinâmicos e de conteúdo variável	23
1.2	Proposta e possíveis aplicações	24
1.3	Roteiro	26
2	<i>Variable Data Printing</i> e documentos adaptativos	27
2.1	Introdução	27
2.2	Conceitos e evolução das tecnologias de VDP	28
2.2.1	<i>Workflow</i> baseado em <i>templates</i>	28
2.2.2	Flexibilização de <i>layouts</i>	30
2.3	Diagramação automática de documentos	30
2.4	Medidas estéticas para diagramação automática de documentos	32
2.5	Considerações sobre o capítulo	33
3	Trabalhos relacionados	35
3.1	Diagramação de documentos com algoritmos aproximativos	35
3.2	Diagramação de documentos baseado em <i>templates</i> e restrições lineares	36
3.3	Diagramação de documentos com otimização combinatória e busca local	36
3.4	Considerações sobre os trabalhos relacionados	37
4	Solução proposta	39
4.1	Diagramação de documentos como um problema de otimização	39
4.2	Modelo proposto	40
4.2.1	Diagramação de páginas de jornal	40
4.2.2	Ferramentas utilizadas	42
4.3	Descrição do algoritmo	44
4.3.1	Visão geral do método	44
4.3.2	Modelagem do método	44
4.3.3	Arquivo de entrada	45
4.3.4	Cálculo de áreas e normalização	46
4.3.5	Estrutura de dados	48
4.3.6	Algoritmo de bissecção	49
4.3.7	Renderização e geração do documento final	52
4.3.8	Preservação da ordem de leitura	54
4.4	Paginação do documento	55
4.5	Melhorando os resultados gerados	58
4.5.1	Medida de erro da página	58
4.5.2	Redimensionando imagens	59

4.5.3	Mantendo elementos agrupados	59
4.6	Objetivos atingidos com o método proposto	61
5	Avaliação empírica e resultados	63
5.1	Avaliação qualitativa	63
5.1.1	Descrição do experimento realizado	63
5.1.2	Descrição da amostra de voluntários	65
5.1.3	Descrição dos resultados	66
5.1.4	Interpretação dos resultados	67
5.2	Testes de performance	68
6	Considerações finais e trabalhos futuros	71
	Referências	73
	Apêndice A Casos de teste aplicados para a avaliação qua- litativa	77
	Apêndice B Formulários para a avaliação qualitativa	99

1 Introdução

1.1 Motivação

Com a convergência dos meios de comunicação digitais (computadores pessoais, PDAs, telefones celulares, entre outros), surge uma demanda por novos mecanismos para composição e intercâmbio de documentos eletrônicos. Documentos são os elementos básicos para troca de informações entre indivíduos e empresas, e existem sob diversas formas: folhetos com ofertas de produtos e serviços, panfletos turísticos, livros, jornais, álbuns de fotografias e extratos bancários, entre outros. As ferramentas para composição de documentos exigidas atualmente devem ser capazes de separar a forma de um documento de seu conteúdo e manipular esta forma, adaptando um mesmo documento para diversos dispositivos diferentes, fazendo com que seu conteúdo possa ser lido e entendido com facilidade.¹ Este tipo de documento é chamado de *Documento Dinâmico* ou *Documento de Conteúdo Variável*. Em paralelo a esta necessidade de adaptação para dispositivos diversos, há uma crescente tendência nas empresas no que diz respeito à personalização do conteúdo que é entregue aos clientes em forma de documentos. Na Subseção 1.1.1 será descrito o processo de personalização de documentos.

1.1.1 Documentos dinâmicos e de conteúdo variável

O termo *Variable Data Printing* (VDP) [1] é usado para designar o processo automático de criação de documentos com conteúdo personalizado. A personalização de documentos permite transmitir ao destinatário documentos únicos, contendo somente as informações relevantes para o mesmo. Este processo também é conhecido como personalização em massa, porque ocorre normalmente em cenários onde devem ser enviados documentos para um grande número de clientes, por exemplo, cartas-padrão (mala direta) de um banco enviadas para cada cliente (com o nome e outras informações do cliente variando de carta para carta) ou folhetos com ofertas de produtos similares aos interesses do destinatário. Este tipo de personalização é bastante usado por empresas em campanhas de *marketing*, visto que um documento personalizado para um determinado cliente (contendo informações a respeito deste cliente em particular) pode ser mais atrativo para que este se interesse pelo anúncio [1].

Os métodos atuais da indústria de impressão para personalização de documentos envolvem o

¹Isto é, considerando as limitações de tamanho, forma e cores do dispositivo em questão.

preenchimento do conteúdo variável em regiões pré-determinadas de um modelo de documento, chamado *template* – normalmente produzido manualmente por um *designer* gráfico –, o que torna o documento inflexível para se adaptar ao conteúdo inserido, caso o *designer* do modelo não tenha conhecimento de como é o conteúdo inserido no mesmo. Além disso, a *Internet* hoje torna possível a obtenção de um grande volume de conteúdo que é atualizado constantemente, com o uso de agregadores de notícias RSS por exemplo [2].

Por isso é também necessária uma forma de adaptar automaticamente a forma de um documento para que este apresente conteúdo de forma flexível, isto é, que o modelo do documento possa ser modificado automaticamente de acordo com a informação que é inserida, sem prejudicar a sua qualidade visual e apresentando todo o conteúdo desejado, garantindo a independência de dispositivos de saída e adaptando-se às necessidades de quem recebe o documento. Adaptar um documento significa criar e/ou organizar os componentes do mesmo (regiões de texto e/ou figuras), de forma que o conteúdo (i. e., informação) do documento final esteja disposto de uma maneira visualmente compreensível e esteticamente agradável para o destinatário deste documento [3, 4]. A Seção 1.2 apresenta a proposta deste trabalho para o problema de diagramação automática de documentos.

1.2 Proposta e possíveis aplicações

Neste trabalho, será apresentado um algoritmo capaz de fazer automaticamente a diagramação de um documento, tendo como entrada um conjunto mínimo de informações, como o conteúdo que deve ser inserido e as dimensões da página (i. e., as dimensões do dispositivo de apresentação do documento).² Como justificativa para este trabalho, é válido dizer que recentemente tem havido um grande interesse por parte de empresas e pesquisadores [2, 5–9] em desenvolver soluções de diagramação automática de documentos, visto a sua importância, em especial para empresas que trabalham na área de impressão e imagens (por exemplo *Kodak* [8], *Xerox* [3, 9] e *HP* [5–7, 10]).

Muitas destas soluções [7–9], como será visto no Capítulo 3, utilizam algoritmos aproximativos (no caso, algoritmos genéticos [11]) para realizar a adaptação automática, na tentativa de obter uma página visualmente agradável através da maximização de qualidades visuais do documento. Um dos problemas desta abordagem é que, no contexto de VDP, o seu uso para produção em massa de documentos se torna inviável devido ao tempo exigido por um algoritmo genético para encontrar uma solução razoável e a imprevisibilidade dos resultados gerados (não-determinismo). Alguns autores tentam simplificar os parâmetros do problema, limitando o espaço de busca através do uso de *grids* e *templates* [12], ou simplificando os parâmetros do documento para o uso de algoritmos determinísticos, como os de programação linear [6, 13], o

²Neste trabalho será usado o termo *diagramação* ou *layout* tanto para descrever o processo de construção de um documento quanto para denotar a organização visual dos elementos de um documento.

que pode tornar difícil a especificação da forma do documento desejado. Soluções puramente determinísticas já existem [5], porém ainda é difícil obter uma solução genérica que integre texto e imagens de forma eficiente [14].

Tendo em vista estes obstáculos, é importante ressaltar que o trabalho aqui proposto ainda não é uma solução genérica para adaptar documentos de qualquer tipo. No Capítulo 4 são apresentadas as características de um problema de *layout* e as diferentes metodologias para se construir uma página. Será visto que a geração de *layouts* completamente livres é um problema complexo em termos computacionais, pois deve ser procurado um *layout* esteticamente agradável em um espaço de busca possivelmente muito grande. Por isso, neste trabalho foi adotada uma estratégia mais simples e eficiente para se construir páginas, seguindo os guias de estilo utilizados em *design* de jornais [15], por ser um modelo de documento familiar e de fácil leitura para a grande maioria das pessoas. Com isto, foi obtido um algoritmo que é capaz de obedecer a diversas restrições impostas por muitas formas de documento, como a preservação da ordem de leitura dos conteúdos, do agrupamento entre elementos relacionados, e de critérios estéticos que afetam significativamente o impacto visual do documento [10] (alinhamento, regularidade, entre outros).³ Mais detalhes sobre o algoritmo desenvolvido são discutidos no decorrer deste texto, em especial no Capítulo 4.

As áreas de aplicação para este trabalho são diversas, visto que a distribuição de conteúdo personalizado está se tornando cada vez mais importante [9]. Além de complementar o processo de *Variable Data Printing* tornando os documentos mais flexíveis em relação ao *layout*, a construção automática de documentos permite a produção de documentos a partir de dados que estão em constante atualização, como o caso de notícias RSS discutido na Subseção 1.1.1, onde um *designer* gráfico humano não teria tempo suficiente para produzir páginas feitas para muitas formas diferentes de conteúdo. Isto cria novos modelos de negócio para empresas especializadas em distribuição de conteúdo e/ou notícias, por exemplo. Até o momento, foram identificados os seguintes serviços – a título de exemplo – que podem fazer uso de tecnologias de *layout* automático, citados a seguir:

- Uma solução para *Web to print* [17], ou seja, a partir da seleção de conteúdo de um *website*, é montado um documento equivalente mas formatado para impressão;
- Construção automática de álbuns de fotos [5];
- Montagem de jornais diários com notícias selecionadas pelo leitor;
- Formatação de documentos com conteúdo agregado da *internet* (i. e., RSS) [2];
- Geração de outros conteúdos, como catálogos e panfletos personalizados, ou seja, atuando na mesma área que os processos de VDP;

³Neste trabalho a ordem de leitura será assumida como sendo a ordem normalmente aplicada em culturas ocidentais, lendo o documento de cima para baixo, e da esquerda para a direita [16].

- Readaptação de um mesmo documento para mídias de tamanhos diferentes.

1.3 Roteiro

Este trabalho está organizado como segue: no Capítulo 2 são apresentados conceitos em relação ao contexto onde este trabalho está inserido, como documentos em *Variable Data Printing*, medidas estéticas de documentos e *layout* de documentos. O Capítulo 3 descreve alguns trabalhos relacionados na área, fazendo um breve comparativo dos objetivos e abordagens utilizadas em relação ao algoritmo aqui apresentado. Após isto, são descritos no Capítulo 4 a modelagem e o funcionamento da nova solução. Os resultados dos testes realizados para a avaliação deste trabalho são apresentados no Capítulo 5. Finalmente, as considerações finais e trabalhos futuros são apresentados no Capítulo 6.

2 *Variable Data Printing* e documentos adaptativos

2.1 Introdução

Na indústria de impressão denomina-se o processo de personalização em massa de documentos como *Variable Data Printing* (VDP). Os documentos utilizados nestes processos, conhecidos também como *Variable Information Documents* (VID) [9], *Documentos Adaptativos* ou *Documentos de Conteúdo Variável* normalmente contêm áreas comuns entre todas as instâncias geradas a partir de um modelo (chamado de *template*, explicado na Seção 2.2) de documento, assim como áreas que são *variáveis*. Contas de telefone, malas-diretas destinadas a diversos destinatários são exemplos simplificados de documentos adaptativos. Estes possuem informações variáveis que são relativas a cada destinatário (por exemplo, número da conta, nome, endereço, entre outros), assim como informações que permanecem iguais em todas as instâncias do documento (por exemplo, logotipo da empresa). Desta forma, a indústria de impressão evoluiu de um modelo em que eram impressas muitas cópias de um único documento para um novo paradigma em que cada documento é único, possui uma única cópia, e é impresso sob demanda [1, 18].

VDP permite que cada instância de um documento seja personalizada de forma que uma campanha de *marketing* direto possa usar dados coletados de várias fontes, a fim de sugerir produtos que são atrativos para o destinatário (isto é, o cliente). A tendência atual nesta área é que através da personalização em massa de documentos seja contornado o desinteresse do cliente, capturando sua atenção com uma propaganda personalizada, impressa unicamente para ele [1].

Este capítulo está estruturado como segue. Na Seção 2.2 será visto como as tecnologias de VDP evoluíram para permitir uma maior flexibilização no *layout* de documentos adaptativos e os desafios atuais. A seguir, na Seção 2.3 serão discutidos brevemente os requisitos necessários para que seja possível automatizar o *layout* de documentos adaptativos para os processos (*workflow*) de VDP. Na Seção 2.4, é introduzido o conceito de estética de documentos, que servirá como base para a construção e análise do método proposto neste trabalho, descrito no Capítulo 4. Finalmente, as considerações sobre os tópicos apresentados neste capítulo são feitas na Seção 2.5.

2.2 Conceitos e evolução das tecnologias de VDP

2.2.1 Workflow baseado em templates

Em aplicações tradicionais de VDP a criação de documentos é feita manualmente por um artista (*designer*) gráfico. O *designer* cria – além do estilo – a diagramação (*layout*) geral do documento, que é preparado para ser apresentado apenas com dimensões fixas. Este *layout* inclui regiões estáticas como texto e imagens, e regiões vazias, conhecidas como *copy-holes*, onde os dados variáveis são inseridos em tempo de impressão. Este modelo de *layout* com regiões vazias é chamado de *template* no contexto de VDP [19]. Com o *template* do documento construído pelo *designer*, são definidas regras sobre como associar o conteúdo – originado de uma base de dados ou outra fonte – com os *copy-holes*. Então, durante a impressão dos documentos, a aplicação de VDP produz automaticamente – através de um processo de *renderização* [20] – instâncias deste *template* através da inserção dos dados variáveis nas respectivas regiões no documento, resultando no documento impresso com dados variáveis. A Figura 1 ilustra este processo (*workflow*) tradicional de VDP.

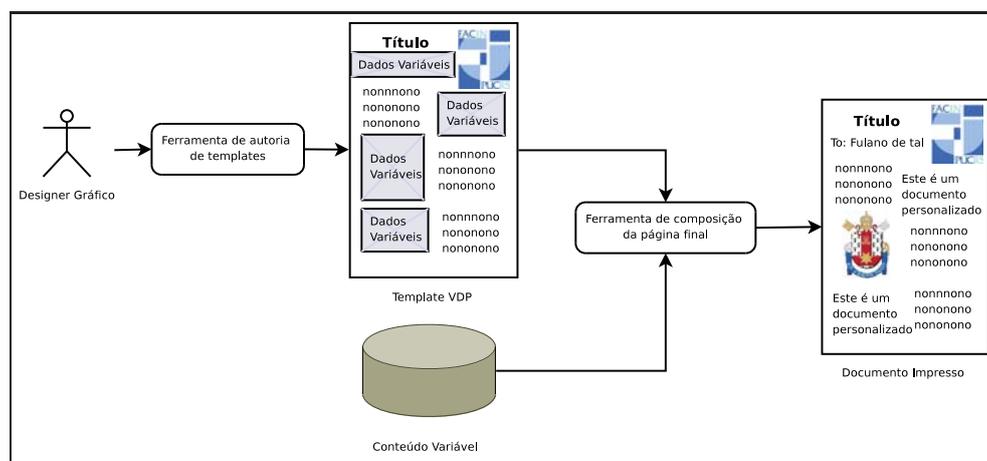


Figura 1: Workflow de um ambiente de VDP.

Os documentos resultantes deste processo são tipicamente bastante similares. Cada *copy-hole* do documento possui um conteúdo associado, que possui um tamanho fixo para ser colocado, e o *layout* geral é o mesmo para todas as instâncias do documento, independente do conteúdo. Este processo tradicional de produção de documentos de conteúdo variável além de exigir o trabalho manual elaborado do *designer* gráfico, também não é capaz de ajustar o documento em função da quantidade ou tamanho do conteúdo relacionado com os *copy-holes*, ou ainda, em função do dispositivo que irá apresentar o documento. Esta forma de *Variable Data Printing*, onde é feito um preenchimento simples dos dados variáveis em cada *copy-hole* em um *template* estático de documento, pode ser considerada como a forma mais simples de se trabalhar com documentos adaptativos. A Figura 2 mostra um exemplo de documentos de conteúdo

variável. O *template* utilizado nos panfletos permanece o mesmo, enquanto que a imagem do carro, as cores, e o texto descritivo variam.

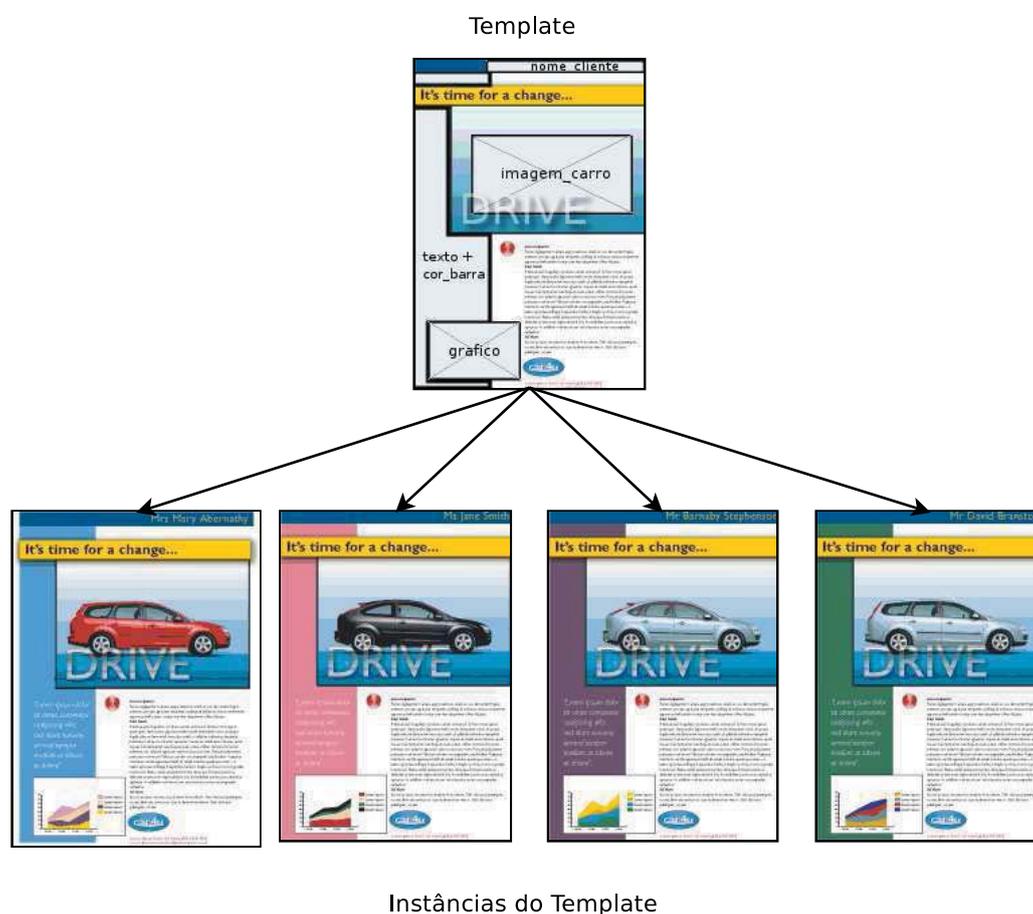


Figura 2: Um exemplo de *template* de documento de conteúdo variável e quatro instâncias geradas a partir do mesmo [21].

É importante salientar que, como a impressão de documentos em VDP é feita sob demanda [1], o processo de geração de uma instância do documento a partir de um *template* e de dados variáveis é executado juntamente com a impressão dos documentos, ou seja, a composição de documentos deve ser um processo eficiente, a fim de criar um grande número de documentos únicos para alimentar a impressão em um tempo viável [22]. Para isto, existem diversas tecnologias [19] para especificação de *templates* de documentos de conteúdo variável. Uma destas tecnologias é a linguagem PPML [18], implementada por impressoras de porte industrial, tais como a *HP Indigo* [1,23], capaz de especificar documentos de conteúdo variável separando seus componentes estáticos e variáveis. PPML é uma das tecnologias mais completas na área de VDP, mas ainda assim continua limitada porque os *templates* usados para especificar o *layout* dos documentos são rígidos em relação aos conteúdos variáveis [19]. Os problemas gerados por essa inflexibilidade serão vistos na Subseção 2.2.2 a seguir.

2.2.2 Flexibilização de *layouts*

Documentos de conteúdo variável estão começando a ser utilizados em situações mais complexas, onde o conteúdo tende a diferir bastante de uma instância do documento para outra. Algumas aplicações exigem uma flexibilidade maior do *template* do documento [22], como por exemplo:

- uma conta telefônica que possui o número de páginas variável para apresentar a lista de ligações feitas pelo usuário;
- catálogos personalizados de produtos, que podem possuir um número variado de conteúdos para inserção no documento;
- relatórios onde os textos e imagens têm tamanhos variados, além do tamanho pré-definido no *template* para os *copy-holes*.

Estas aplicações requerem que o *layout* do documento resultante não seja um *template* rígido, mas sim que seja capaz de se adaptar ao conteúdo, seja em formato, seja em número de páginas. Na Seção 2.3 será apresentado o conceito de construção automática de *layouts* e os problemas envolvidos neste processo.

2.3 Diagramação automática de documentos

A tecnologia atualmente exigida por aplicações de VDP é um mecanismo de *layout* automático, cuja função é ajustar o *layout* original ou criar um *layout* novo para apresentar os conteúdos variáveis [3, 24]. Porém, construção automática de *layouts* é um problema complexo, porque envolve otimização de posições e áreas na geometria da página [24]. A intenção do trabalho aqui proposto é o de flexibilizar o *layout* dos documentos, observando os seguintes aspectos:

- O *layout* do documento deve ser adaptado ao conteúdo variável, e não fixado por um *template* rígido, garantindo que o documento possa atender a conteúdos mais diversificados. O uso de *templates*, apesar de ser um mecanismo controlado e eficiente, adiciona complexidade no processo de impressão e não podem ser generalizados para conteúdos variados [25];
- Deve ser permitido que um mesmo conjunto de conteúdos variáveis possam ser impressos em páginas de tamanhos diferentes, garantindo assim a adaptação de um mesmo documento para mídias diversas;
- O *layout* deve ser construído rapidamente, a fim de não congestionar o processo de VDP;

- A intervenção humana no processo (como a necessidade de um *designer* gráfico para readaptar um *template* para conteúdos variados) deve ser reduzida ao máximo, já que a produção de documentos em VDP deve ser feita em massa e com conteúdos dinâmicos [2].

Na Figura 3 são mostrados dois documentos simulando capas de jornal, cujo *layout* foi gerado de forma automática pelo algoritmo desenvolvido neste trabalho, a partir do conteúdo informado. O documento à direita, na Figura 3(b), mostra como o *layout* foi readaptado após a inserção de uma imagem no documento à esquerda, na Figura 3(a).



(a) Uma capa de jornal gerada automaticamente.

(b) Readaptação da capa gerada na Figura 3(a).

Figura 3: Um exemplo de documento cujo *layout* foi gerado automaticamente a partir do conteúdo, e a readaptação do mesmo documento após a inserção de uma nova imagem como conteúdo.

O uso de um processo de construção automática de documentos – descrito em detalhes no Capítulo 4 –, apesar de compensar a inflexibilidade do uso de *templates*, possui algumas desvantagens:

- O usuário tem pouco controle sobre o layout gerado;
- Um *layout* gerado para um conjunto de dados não pode ser reaproveitado para conteúdos diferentes, sendo necessário portanto a reconstrução do *layout* quando o conteúdo varia [24];
- Devem ser observados um conjunto de critérios estéticos para a produção do layout [3].

Para que um sistema seja capaz de produzir *layouts* automaticamente, é necessário que este saiba como dispor os conteúdos na página de uma forma coerente para o leitor do documento final. Para isso, na Seção 2.4 a seguir serão vistas algumas medidas estéticas para documentos, que servem como guia para o processo de construção de *layouts* de páginas.

2.4 Medidas estéticas para diagramação automática de documentos

Estética é considerado um dos atributos mais importantes na organização de um documento. Guias para se produzir documentos com uma aparência agradável geralmente são encontradas em livros de *design* gráfico [4, 26], que normalmente transmitem técnicas para construção de documentos através de exemplos, mostrando práticas boas e ruins para a manutenção estética de um documento. Estas métricas permitem um grau razoável de flexibilidade de estilo e opções de *layout* que podem auxiliar na classificação de um documento [3].

Crítérios para se medir estética de documentos adaptativos são importantes por diversos motivos na área de VDP. Primeiro, documentos de conteúdo variável são produzidos em grandes quantidades, e com isso podem sofrer diversas transformações durante o seu processo de construção. Por isso, é necessária uma maneira de se medir o quão distante uma instância de documento está em relação ao modelo original, em termos de estética [10]. Segundo porque com as necessidades de um *layout* gerado de forma automática, estes critérios servem como guia para que o algoritmo de diagramação automática construa *layouts* que obedeçam às noções de qualidade conhecidas em *design* gráfico [4].

Existem esforços por parte de pesquisadores [3,10,27,28] em quantificar estas qualidades de um documento, através de funções matemáticas que medem características de um documento associadas às qualidades sugeridas por artistas gráficos. No trabalho de Harrington et al. [3] são descritos diversos critérios estéticos para avaliar um documento. No que diz respeito a *design* gráfico, as noções de estética tendem a ser subjetivas, mas Harrington et al. [3] propuseram medidas de qualidade de um documento através de modelos matemáticos que tentam mapear as noções de estética de documentos encontradas em livros de *design* e projeto gráfico [4]. A seguir são descritos alguns destes critérios, considerados mais relevantes para este trabalho:

- **Critério de alinhamento:** indica que as regiões de um documento devem estar em um padrão alinhado, ou seja, as bordas das regiões de texto ou figuras interceptam uma mesma linha ou coluna, sejam elas centralizadas, na esquerda ou na direita;
- **Critério de regularidade:** indica que as regiões tendem a repetir em certos pontos de forma mais previsível, garantindo que estas tenham alturas e larguras parecidas. Isto também é conhecido como *Repetição* [4], um objetivo estético mais geral;

- **Critério de uniformidade:** indica que as regiões são distribuídas de forma homogênea pela página, ou seja, a área da página é bem distribuída para as regiões que a ocupam;
- **Critério de fluxo de espaço em branco:** indica que todo o espaço em branco contido na página está conectado às margens da página, ou seja, não há nenhuma região vazia “presa” dentro da página do documento;
- **Critério de proporção:** indica que o *aspecto* (a razão entre largura e altura) de cada região tende a ser parecido com um aspecto pré-determinado, considerado “agradável” para leitura (por exemplo, uma região quadrada tem aspecto de valor 1.0).

Outros trabalhos [5,7,8] propuseram critérios similares para a classificação do *layout* gerado pelos algoritmos, e é mostrado no Capítulo 4 de que forma os critérios citados anteriormente são preservados ou aproximados pelo algoritmo proposto neste trabalho. Apesar de não ser obtido sempre o melhor *layout* possível (este conceito é definido no Capítulo 4) para um dado conteúdo de entrada, por razões de decisão de projeto, optou-se por manter o algoritmo simples, que, ainda assim é capaz de gerar documentos que mantêm um bom equilíbrio entre os critérios estéticos citados.

Um outro critério que pode ser considerado importante na avaliação de qualidade de um documento, é a preservação da ordem cronológica de leitura dos elementos que compõem o documento ou a proximidade [4] entre elementos que possuem relações entre si (isto é, se os conteúdos são relacionados).

2.5 Considerações sobre o capítulo

Este capítulo dedicou-se a apresentar conceitos gerais que estão no contexto do trabalho proposto. A seguir, no Capítulo 3, são descritos os principais trabalhos na área que propuseram soluções para o problema de *layout* de documentos.

Os requisitos impostos pela área de VDP descritos neste capítulo servirão como guia para a análise crítica dos trabalhos relacionados no Capítulo 3 e usados como objetivos para o algoritmo proposto no Capítulo 4.

3 Trabalhos relacionados

Este capítulo apresenta os trabalhos relacionados já publicados na área de *layout* de documentos. O foco deste capítulo será apenas nos trabalhos mais relevantes que tratam o problema específico de produção automática de documentos, deixando de lado a discussão de técnicas de *layout* geométrico em geral, tais como *Tree Maps* [29], *Bin Packing* [30, 31], ou *layout* de grafos [32, 33]. Esta decisão foi tomada tendo em vista que estes algoritmos não consideram restrições necessárias em *layout* de documentos, como preservação de ordem de leitura, produção em massa (i. e., eficiência), entre outros.

Nas seções a seguir é apresentado resumidamente o estado da arte na área de *layout* de documentos, organizado pela estratégia utilizada pelo autor para abordar o problema. A Seção 3.1 descreve trabalhos que fazem uso de métodos estocásticos de otimização, tais como algoritmos genéticos e *simulated annealing* [11, 34, 35]. Já na Seção 3.2 é descrito um método em que o problema de *layout* é tratado como um sistema de restrições e é resolvido através de programação linear [13]. Algumas estratégias determinísticas para *layout* são descritas na Seção 3.3, envolvendo o uso de programação dinâmica e busca local [13]. Finalmente, na Seção 3.4 são apresentadas algumas considerações sobre os trabalhos relacionados aqui descritos.

3.1 Diagramação de documentos com algoritmos aproximativos

Goldenberg [7] propôs um algoritmo genético para resolver o problema através de recortes sucessivos em uma página. Porém, além dos problemas inerentes a algoritmos genéticos (não-determinismo, baixa performance), os *layouts* gerados não eram necessariamente agradáveis, pois, da forma como o algoritmo foi projetado, critérios estéticos como alinhamento e homogeneidade eram dificilmente preservados pela função de *fitness* [11] utilizada.

Purvis et al. [9] propuseram uma solução diferente utilizando algoritmos genéticos, em que os conteúdos (texto, imagens, entre outros) podiam assumir formas livres, e o espaço de busca era construído variando-se as dimensões e posição de cada conteúdo. A função de *fitness* era definida como uma composição de critérios estéticos, como propostos por Harrington et al. [3], que funcionavam como um conjunto de restrições para as qualidades desejadas para a página [34]. Porém, como apontado pela autora, o algoritmo genético dificilmente convergia para uma solução satisfatória, o que exigia execuções múltiplas do algoritmo a fim de obter um resultado agradável, o que não era garantido sempre. Geigel e Loui [8] propuseram um

algoritmo genético similar (com critérios estéticos também similares) para gerar álbuns de fotografia, porém introduziram conceitos como ordenação de conteúdo (cronologia, apesar de não ser garantida sempre) e criação do documento possuindo mais de uma página.

Johari et al. [36] propuseram um algoritmo de *layout* específico para construir páginas amarelas de listas telefônicas, utilizando métodos aproximativos, no caso *Simulated Annealing* [35]. Em páginas amarelas devem ser colocados anúncios, que são retângulos de aspecto e área fixas, e textos, que podem sofrer quebras de linha e percorrer a área livre da página. Esta combinação de formas rígidas e flexíveis torna o espaço de busca do problema bastante complexo, pois dois aspectos devem ser otimizados simultaneamente: a colocação dos anúncios, similar a um problema de *bin packing* [30], e a inserção de textos na área livre da página, que deve ser aproveitada da melhor maneira possível.

3.2 Diagramação de documentos baseado em *templates* e restrições lineares

Lin [6, 24] descreve uma abordagem que foge do uso de algoritmos genéticos, transformando os critérios desejados (como posicionamento e aspecto) do conteúdo em restrições lineares, transformando o problema em um problema de programação linear, podendo ser resolvido por algoritmos determinísticos, como o *simplex* [13]. O problema desta abordagem é que, além do seu baixo desempenho, proibitivo para aplicações de VDP, um *template* também deve ser especificado, através de restrições lineares com um nível de abstração muito baixo (ex. $elemento1.altura > 10$, $elemento1.x < elemento2.x$) para ser especificado por um usuário comum, preocupado com restrições de alto nível (ex. *conteúdo deve estar alinhado à esquerda*). A abordagem descrita aqui tenta exigir o mínimo de entrada do usuário (o conteúdo a ser alocado, e parâmetros relacionados com a página onde serão alocados os conteúdos), mesmo que isso signifique uma menor liberdade na forma do documento.

3.3 Diagramação de documentos com otimização combinatória e busca local

Atkins [5] propôs um algoritmo determinístico para produzir *layouts* rápidos de pôsteres de fotografia. Porém, além de trabalhar apenas com imagens, que possuem aspecto fixo (ao contrário de texto, cujo aspecto permite variações), o algoritmo não é capaz de manter as fotografias na ordem em que elas são dadas pelo usuário, o que não garante a ordem de leitura da página.

Recentemente, de Oliveira [14] propôs dois algoritmos determinísticos capazes de produzir *layouts* de documentos, cada um destes voltado para uma aplicação específica. Um dos algo-

ritmos recebia o conteúdo em forma de retângulos com geometria fixa (largura e altura) para serem dispostos em uma página de documento sem sofrer alterações de aspecto e área, e outro permitia formas livres para os elementos, que possuíam apenas a informação de área associada. O algoritmo então, através do uso de programação dinâmica [13] encontrava soluções ótimas baseado em uma função de medida de erro em uma página. Porém, apesar dos bons resultados gerados, do tempo de execução dos algoritmos, como apontado pelo autor, é proibitivo para processos em *Variable Data Printing*.

3.4 Considerações sobre os trabalhos relacionados

Após um estudo sobre os principais trabalhos existentes em diagramação de documentos até o momento deste texto, chega-se à conclusão de que um algoritmo que gere *layouts* mais simplificados, mas que seja capaz de gerar um documento bem estruturado (também em termos estéticos) e com um grau maior de previsibilidade – tanto no tempo de execução quanto nos documentos gerados – pode ser mais interessante para aplicações em VDP, que exigem justamente este tipo de confiabilidade na hora de produzir documentos personalizados em massa. Além disso, é também de interesse que a interação entre o algoritmo de *layout* e o usuário seja mínima, o que facilita a produção automatizada de documentos. Isto pode ser útil em cenários além de VDP, onde conteúdos são atualizados constantemente, e um *layout* para adaptar este novo conteúdo deve ser gerado dinamicamente, à medida em que os conteúdos são alterados. Um exemplo de aplicação para isto seria a geração de documentos contendo notícias obtidas através de *web feeds*. Schrier et al. [2] propuseram uma solução para esta aplicação, mas ainda exige o uso de *templates* construídos manualmente, o que não é desejado neste trabalho.

A solução proposta neste trabalho será discutida a seguir, de maneira detalhada, no Capítulo 4.

4 Solução proposta

Este capítulo apresenta a solução proposta neste trabalho para diagramação automática de documentos. O roteiro deste capítulo é dado como segue: na Seção 4.1 o algoritmo proposto neste trabalho é categorizado em relação ao problema mais geral de *layout*, que é um problema de otimização [37]. Na Seção 4.2 é apresentado o modelo utilizado para validação deste trabalho para o qual este algoritmo produz *layouts*, que são documentos similares a capas de jornal. Neste modelo também são apresentadas as ferramentas utilizadas no processo de geração de páginas, que fazem a interface entre a entrada de dados e a saída final do documento. Na Seção 4.3 o algoritmo básico é descrito de forma detalhada. A seguir, na Seção 4.4 é apresentado o algoritmo complementar capaz de dividir os elementos de um documento em um *layout* composto por duas ou mais páginas. Na Seção 4.5 são mostradas técnicas e medidas para melhorar o algoritmo básico descrito na Seção 4.3. Finalmente, na Seção 4.6 são resumidos os objetivos atingidos sobre o método proposto.

4.1 Diagramação de documentos como um problema de otimização

Como discutido na Seção 1.2, o método aqui proposto não é uma solução genérica para distribuir conteúdos em um documento. Isto porque o espaço de possibilidades para se construir um *layout* livre de restrições é muito grande [9, 14, 36], visto que parte do problema de distribuir elementos em uma página pode ser reduzido – a grosso modo – a um dos seguintes problemas:

- Empacotamento, ou *Bin Packing* [30], caso os elementos sejam retângulos fixos;
- *Cutting Stock* ou *Tree Maps* [7, 29], caso os elementos tenham aspecto livre. Neste caso deve ser buscada a melhor maneira de se dividir uma página em diversas partes para cada elemento.

A redução do problema de *layout* automático para um destes dois problemas caracteriza o problema de *layout* automático como um problema de otimização combinatória [37], cuja solução ótima possuiria a mesma complexidade assintótica destes dois problemas, conhecidos por pertencerem à classe de complexidade *NP-Difícil* [29, 30].

Como visto no Capítulo 3, a maioria dos autores tentou compensar a complexidade do problema limitando o seu espaço de busca, seja com o uso de *templates* de documentos [12, 24]

ou impedindo a explosão combinatória do problema, através da fixação de ordem dos elementos por exemplo [36]. Neste trabalho esta redução de problema também ocorreu, mas foram escolhidas restrições que beneficiam o *layout* do documento, como será visto no decorrer deste capítulo.

Além disso, o algoritmo apresentado neste trabalho não encontra soluções ótimas para o modelo proposto, ou seja, as páginas produzidas pelo algoritmo não são necessariamente as melhores soluções considerando critérios de ocupação da página pelo conteúdo e critérios estéticos. Porém, será visto no decorrer deste capítulo que a heurística desenvolvida neste trabalho produz resultados bastante satisfatórios, levando em consideração os requisitos da área de VDP, discutidos no Capítulo 2, ao contrário de métodos aproximativos para otimização [9] como proposto em outros trabalhos.

4.2 Modelo proposto

Esta seção descreve o modelo que servirá como embasamento para o algoritmo desenvolvido neste trabalho, definindo a classe de diagramações produzidas, e também utilizado para a construção dos casos de teste utilizados na avaliação qualitativa, como descrito no Capítulo 5. Este modelo é descrito a seguir na Subseção 4.2.1.

4.2.1 Diagramação de páginas de jornal

O trabalho desenvolvido foca em documentos que possuem um *layout* modular (hierárquico), dividido em regiões retangulares e alinhadas em colunas, como é feito em documentos como páginas de revistas e/ou jornais.

Um modelo de documento como jornais é interessante para se trabalhar, pois os critérios estéticos são simples de definir [15], e produz *layouts* de fácil leitura, e que a grande maioria de pessoas já está familiarizada, evitando surpresas ou dificuldades de leitura.

Mais especificamente, os principais componentes necessários para compor o *layout* de um jornal [15] são de razoável interesse para os processos utilizados em VDP (Capítulo 2), descritos a seguir:

- **Paginação:** é o mecanismo capaz de distribuir uma série de elementos como texto ou imagens em um documento composto por diversas páginas. No caso de jornais, o número de páginas é normalmente um múltiplo de 4, pois cada folha no estilo *Broadsheet* [15] contém textos (matérias) para 4 páginas. O algoritmo para distribuir elementos em páginas é descrito na Seção 4.4;

- **Mistura de regiões de imagem e texto:** imagens em jornais são por natureza elementos retangulares e rígidos, que não podem mudar de aspecto. Textos, em contrapartida, podem fluir pelo espaço livre da página. Um mecanismo automático que combine estes dois tipos de elementos em um único documento ainda é difícil de se obter [22,36]. Porém, no decorrer deste capítulo (em especial na Seção 4.3) será descrito como o método contorna este problema parcialmente;
- **Ordem de leitura, proximidade e alinhamento de textos:** são três critérios essenciais em jornais [15]. A ordem de leitura garante que os textos mais “importantes” da página serão lidos primeiro (a Seção 4.3 mostra como a ordem de leitura é preservada pelo algoritmo desenvolvido). O alinhamento em colunas existente nos jornais ajuda os olhos do leitor a separar textos diferentes [15]. Já a proximidade especifica que elementos relacionados entre si devem aparecer próximos na página, a fim de não confundir o leitor [4], tais como textos associados com imagens em um jornal;
- **Homogeneidade da página:** É um critério essencial em várias formas de documentos [3, 14], que especifica que os elementos devem aproveitar a área da página de forma equilibrada, ou homogênea, preenchendo a página por inteiro, como é o caso em jornais.

O modelo de página utilizado neste modelo é mostrado na Figura 4. Note que isto difere do conceito de *template* apresentado no Capítulo 2 porque não há especificação de regiões para conteúdo variável neste modelo, mas sim uma região onde o *layout* da página será construído.

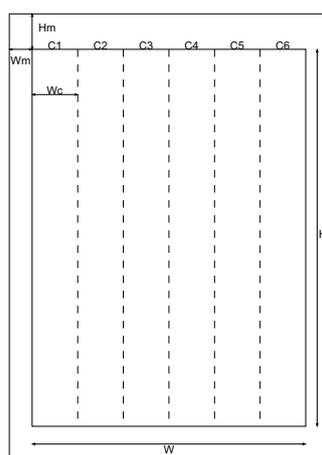


Figura 4: Componentes de uma página de documento utilizado como entrada para o algoritmo, com 6 colunas neste exemplo.

Os parâmetros de um página de documento ilustrados na Figura 4 são os seguintes:

- W, H : a largura e altura respectivamente da região da página onde o conteúdo é distribuído, especificados pelo usuário como entrada para o algoritmo;

- W_m, H_m : os tamanhos das margens horizontais e verticais, respectivamente. Para fins de simplificação, será assumido neste trabalho que ambas as margens à esquerda e à direita na página possuem a mesma dimensão W_m , assim como as margens superiores e inferiores possuem a mesma dimensão H_m . Também não é reservado nenhum espaço para cabeçalhos e rodapés;
- C_1, \dots, C_{NC} e W_c : O número NC de colunas C_1, \dots, C_{NC} que dividem a página verticalmente é especificado pelo usuário como entrada. No caso da Figura 4, $NC = 6$. Além disso, todas as colunas possuem a mesma largura W_c , que pode ser calculada como

$$W_c = \frac{W}{NC}. \quad (4.1)$$

Novamente, é importante ressaltar o fato de que a existência de margens e colunas nas páginas de um documento gerado pelo algoritmo aqui proposto não é o mesmo que um *template* produzido por um *designer* gráfico, e sim apenas uma restrição e simplificação no funcionamento do algoritmo para que este sempre produza páginas alinhadas verticalmente.

A seguir serão descritas brevemente algumas ferramentas e o processo utilizado na implementação deste trabalho.

4.2.2 Ferramentas utilizadas

Durante o desenvolvimento do protótipo da ferramenta para este trabalho, o algoritmo de *layout* foi inserido em um processo (ou *workflow*) similar ao de VDP (Capítulo 2), mas para fins de validação do trabalho, foi permitido ao usuário selecionar o conteúdo através de alguma interface gráfica (não definida neste trabalho, ver Figura 5), capaz de produzir um arquivo XML [38]. Este arquivo especifica o conteúdo que será usado pelo algoritmo de adaptação para produzir o *layout*, e por sua vez, o algoritmo de adaptação produz um arquivo \LaTeX [39] para ser então processado (descrito a seguir) e transformado em um documento PDF.¹ O uso de um formato XML como entrada foi escolhido devido à facilidade de se intercambiar este tipo de formato atualmente [38]. Já o uso do processador de textos \LaTeX foi escolhido para que não houvesse preocupação com problemas de formatação de baixo nível durante o desenvolvimento deste trabalho, tais como quebras de linhas, hifenização, etc. como feito em outros trabalhos mais abrangentes [2, 12, 24]. Mais especificamente, o \LaTeX é capaz de produzir documentos divididos em colunas, e ajustar automaticamente o texto nas regiões produzidas pelo algoritmo de *layout* caso não haja um ajuste exato, como será explicado no decorrer deste capítulo, na Seção 4.3. A Figura 5 ilustra em alto nível o processo de geração de documentos com o algoritmo proposto, e a Figura 6 mostra um exemplo dos arquivos gerados e de como a entrada é transformada em um documento final no *workflow* descrito.

¹Não será discutido neste trabalho o funcionamento destas ferramentas.

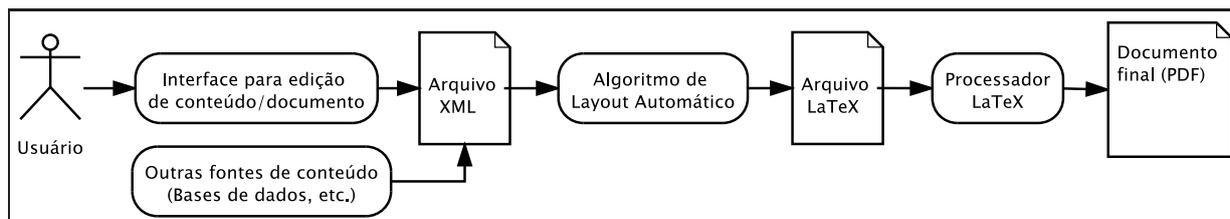


Figura 5: Ilustração do *workflow* utilizado para validação do algoritmo desenvolvido.



Figura 6: Exemplo dos arquivos gerados no processo de geração do documento final.

Uso de outras ferramentas

Cabe ressaltar que de nenhuma forma o algoritmo proposto neste trabalho está atrelado a estas tecnologias, tendo estas sido escolhidos apenas pela sua conveniência e para fins de desenvolvimento do protótipo da ferramenta utilizada para validação e testes deste trabalho. Porém, a ferramenta que realiza a etapa de renderização (descrita na Subseção 4.3.7) deve atender às seguintes capacidades mínimas:

- Ser capaz de manipular imagens e textos;
- Ser capaz de distribuir e formatar texto automaticamente em uma região retangular;
- Ser capaz de adaptar os atributos do texto de forma a preencher plenamente a região que o contém;

- Ser capaz de dividir uma página em duas ou mais colunas verticais.

Até o momento de publicação deste trabalho, apenas três ferramentas são conhecidas pelo autor capazes de atender a estes requisitos: o processador de textos \LaTeX [39], o sistema de editoração eletrônica *Scribus* [40] e a linguagem de formatação *XSL:FO* [20], sendo esta última até o momento incapaz de trabalhar com uma página com textos que ocupam diversas colunas.

4.3 Descrição do algoritmo

4.3.1 Visão geral do método

O algoritmo desenvolvido resolve o problema de alocação de elementos (conteúdo) em uma página através de uma estratégia de *divisão e conquista* [13]. A página do documento passa por um processo denominado de *bissecção*, ou partições recursivas, que divide a página em duas regiões retangulares distintas, através de uma linha horizontal ou vertical que passa de uma borda da página para outra, sempre paralela a um dos eixos da página. Cada região resultante da bissecção é então recursivamente particionada, e este processo termina quando o número de regiões na página é igual ao número de elementos que devem ser alocados no documento. A decisão de onde estas partições ocorrem na página são decididas segundo os critérios do algoritmo de bissecção, que será descrito na Subseção 4.3.6. Ainda, as seguintes restrições se aplicam:

- As bissecções feitas em regiões da página do documento são linhas horizontais ou verticais apenas, paralelas ao retângulo que define a página;
- Como a página de um documento é organizada em um número NC de colunas verticais de tamanhos iguais, isto significa que uma bissecção vertical só poderá ser feita em pontos discretos na linha horizontal da página, garantindo assim o critério de alinhamento (vertical) entre todos os elementos no documento final;
- Em decorrência disto, todos os elementos da página ocupam um número inteiro de colunas, e não podem sobrepor-se uns aos outros.

A seguir será vista a modelagem do problema para a definição do algoritmo.

4.3.2 Modelagem do método

Para abordar o problema e prover uma descrição detalhada e correta do algoritmo no decorrer deste capítulo, as diversas etapas do algoritmo são ilustradas em detalhes na Figura 7.

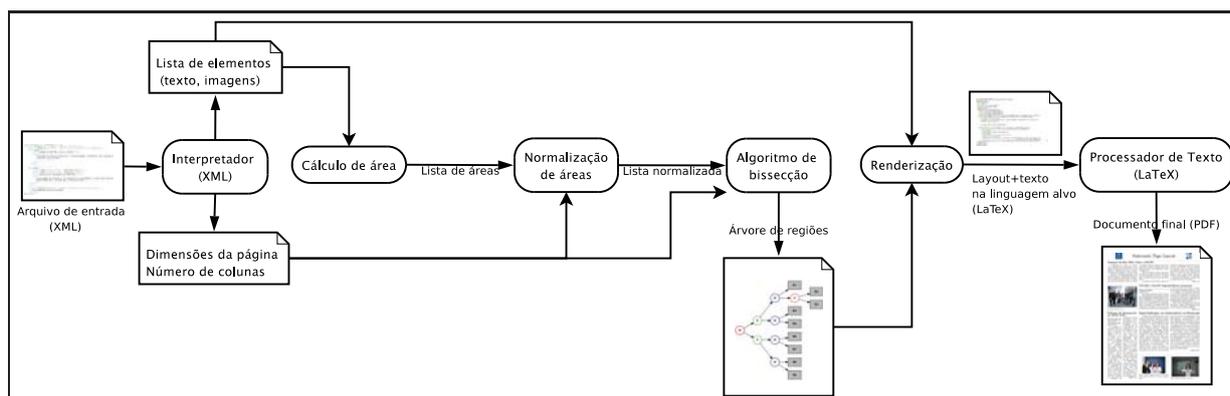


Figura 7: *Workflow* detalhado do algoritmo desenvolvido.

Estas etapas serão descritas a seguir, nas Subseções 4.3.3, 4.3.4, 4.3.5, 4.3.6 e 4.3.7, e podem ser resumidas da seguinte forma:

1. O arquivo de entrada é processado, e são extraídas a seqüência de elementos, sendo estes texto e/ou imagens, e as dimensões e número de colunas da(s) página(s) do documento;
2. Para cada elemento da seqüência recebida é calculada a área estimada correspondente ao tipo de elemento (texto ou imagem);
3. Estas áreas são normalizadas a fim de representar sua quantidade através de proporções do tamanho da página do documento;
4. A região principal (a página) do documento é dividida recursivamente pelo algoritmo de bissecção, alocando cada elemento da seqüência para uma região exclusiva, de acordo com a sua área requisitada. Ao final do processo, o número de subregiões na página é igual ao número de elementos;
5. O processo de bissecção produz uma estrutura de dados como saída, que provê informações sobre as dimensões e posições de cada região produzida e associa cada região com um elemento;
6. O processo de *renderização* é realizado sobre esta estrutura de dados para um formato de saída específico (no modelo proposto neste trabalho, o \LaTeX [39]);
7. O processador de textos específico lê o arquivo de saída produzido pela etapa de *renderização* e produz o documento final.

4.3.3 Arquivo de entrada

O arquivo de entrada para o algoritmo deve conter as seguintes informações:

- As dimensões W (largura), H (altura) e o número NC de colunas verticais relativos à(s) página(s) do documento. Todas as colunas possuirão largura W_c (Equação (4.1));
- A entrada do algoritmo é uma seqüência ordenada (E_1, E_2, \dots, E_n) de *elementos*, que podem ser ou *textos* (i. e., *strings*, uma seqüência de caracteres), ou *imagens* (o endereço de uma imagem no sistema de arquivos). A ordem da seqüência de entrada determina a ordem de leitura dos elementos no documento (segundo o critério da ordem de leitura feita no ocidente [16], mais detalhes na Seção 4.3.8). Junto com cada imagem, deve ser fornecido também o número desejado de colunas que esta imagem deseja ocupar. Isto será discutido na Subseção 4.3.4 a seguir.

4.3.4 Cálculo de áreas e normalização

Para que o algoritmo produza resultados satisfatórios, é necessário que textos compridos exijam áreas maiores da página do documento e textos curtos exijam áreas menores, assim como imagens. Com uma informação quantitativa de quanto espaço um determinado elemento requer em um página, o algoritmo de bissecção é capaz de decidir posições boas para particionar a página.

Para que o usuário não fique encarregado de fornecer estes dados juntamente com a seqüência de elementos, é descrita aqui uma maneira de estimar as áreas requisitadas por textos e imagens.

Estimando a área de textos

Um *texto* é um tipo de elemento que não possui aspecto (ou geometria) fixo, e portanto é definido apenas pela sua área estimada QA_i . A área ocupada por um texto i pode ser calculada de forma aproximada como

$$QA_i = \alpha \times W \times H \times \text{strlength}(i). \quad (4.2)$$

Na Equação (4.2), $\text{strlength}(i)$ é o número de caracteres contido no texto i e α é uma constante que estima o tamanho médio e um único caractere no documento final, relativo à área da página em questão.

Estimando a área de imagens

Uma imagem, diferentemente de um texto, é uma região retangular que possui sua geometria fixa. Para cada imagem i da seqüência de elementos, deve existir um número de colunas C_i fornecido como entrada. Este número define quantas colunas inteiras da página serão ocupadas

pela imagem. Além disso, uma imagem também possui uma *razão de aspecto*, que define a geometria da mesma, definida pela razão

$$AR_i = GW_i/GH_i, \quad (4.3)$$

onde GW_i e GH_i representam as dimensões originais da imagem. Então, se temos o aspecto AR_i da imagem, a largura e altura ocupadas pela mesma na página são

$$W_i = W \times (C_i/N_c) \quad \text{e} \quad H_i = W_i/AR_i \quad (4.4)$$

e a área ocupada pela imagem i é definida como

$$QA_i = W_i^2/AR_i. \quad (4.5)$$

Normalização

Para que estas áreas estimadas para os elementos de um documento sejam úteis para o algoritmo, é necessário que as mesmas sejam *normalizadas* em relação à área da página deste documento. O processo de normalização transforma as áreas requisitadas em frações de uma unidade de página, tornando-as áreas relativas às dimensões da página, ou seja, a soma das áreas relativas A_i de cada elemento i representa 100% da área da página, garantindo assim a homogeneidade e ocupação total da página:

$$\sum_{i=1}^n A_i = 1. \quad (4.6)$$

A área relativa de uma imagem i , devido à necessidade de receber uma área que corresponda exatamente ao número de colunas requerido, pode ser calculada como descreve a Equação (4.7):

$$A_i = QA_i/(W \times H). \quad (4.7)$$

Ou seja, como uma imagem possui geometria fixa e ocupa um número inteiro de colunas da página, basta que esta consuma a área da página de forma proporcional.

No caso de textos, que podem fluir livremente pela página, com áreas menores ou maiores que sua área requisitada, devem distribuir a área restante da página (descontando as áreas requisitadas por imagens) entre si. Assim sendo, a área normalizada de um texto é calculada como

$$A_{img} = \sum_{imagem(i)} QA_i \quad (4.8)$$

$$A_{txt} = \sum_{texto(i)} QA_i - A_{img} \quad (4.9)$$

$$A_i = (W \times H - A_{img}) / A_{txt} \times QA_i, \quad (4.10)$$

onde $imagem(i)$ indica que o elemento i é uma imagem e $texto(i)$ indica que i é um texto.

Após a etapa de cálculo de áreas e normalização, o algoritmo é capaz de particionar a página respeitando as áreas requisitadas por cada elemento. A seguir, na Subseção 4.3.5 será apresentada a estrutura de dados utilizada pelo algoritmo.

4.3.5 Estrutura de dados

Como ilustrado na Figura 7, a execução do algoritmo produz uma estrutura de dados como saída para o renderizador, que deve ser descrita antes de apresentar o algoritmo em si.

Sendo um algoritmo de natureza recursiva, é necessário que haja uma maneira de representar e armazenar as diversas regiões produzidas pelo processo de bissecção. Para isso, a estrutura de dados mais apropriada é uma árvore binária, onde:

- a raiz representa a região “principal”, ou uma página inteira do documento;
- os nós internos representam sub-regiões que são recursivamente divididas em duas;
- as folhas, que são regiões indivisíveis, onde cada uma contém um elemento do documento.

Para representar uma região r , é necessário que cada nó possua os seguintes atributos:

- $r.c$: número de colunas contidas na região;
- $r.x$ e $r.y$: posição absoluta da região;
- $r.w$ e $r.h$: largura (pode ser deduzida a partir de $r.c$ e a largura W da página) e altura da região, respectivamente;
- $r.i$ e $r.j$: intervalo $[i \dots j]$ da seqüência de elementos contida dentro desta região. Se $r.i = r.j$, então o nó não será mais dividido, pois contém apenas um elemento;
- $r.esq$ e $r.dir$: apontadores para as regiões filhas, resultantes de uma subdivisão.

A Figura 8 ilustra como esta estrutura representa as bissecções feitas em uma página e um intervalo de elementos.

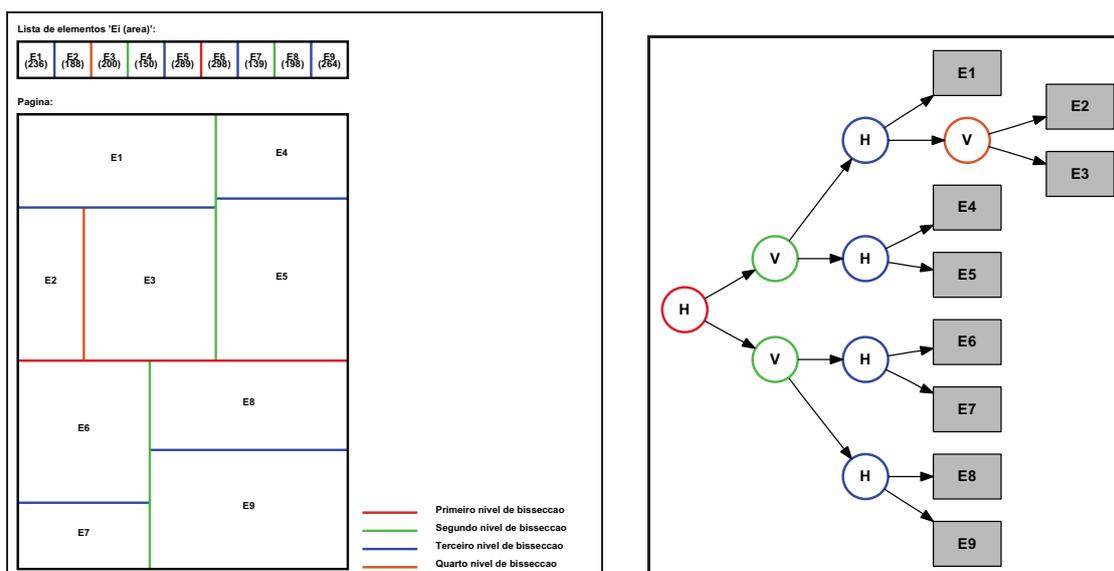


Figura 8: Ilustração do algoritmo e a árvore gerada pelo processo de bissecção.

4.3.6 Algoritmo de bissecção

Como visto no início desta seção, a entrada do algoritmo é uma seqüência (E_1, E_2, \dots, E_n) de elementos. Para cada elemento $E_{1 \leq i \leq n}$, é associada a área $A_i \in \mathfrak{R}$, obtida na etapa de cálculo de áreas e normalização (Subseção 4.3.4). Inicialmente, a página (a região raiz, como descrito na Subseção 4.3.5) contém todo o intervalo de elementos $[1 \dots n]$. Ou seja, para cada região é associado um intervalo $[i \dots j]$, que representa a subseqüência de elementos que está contida naquela região. A cada bissecção, é escolhido um ponto de corte k no intervalo, que divide a seqüência de elementos em duas seqüências, $[i \dots k]$ e $[k + 1 \dots j]$.

O critério de escolha do ponto de corte k na seqüência é determinado pelo ponto *mediano* [41], ou seja, a área total dos elementos $\sum_{l=i}^j A_l$ é dividida no ponto em que a soma das áreas $\sum_{l=i}^k A_l$ é o mais próximo da soma $\sum_{l=k+1}^j A_l$. O efeito desta escolha é a preservação da medida estética de uniformidade (Seção 2.4), pois a área da página é dividida de forma proporcional a cada bissecção para as áreas de cada lado do corte, mantendo homogênea a distribuição de áreas na página.

Para localizar o ponto k na seqüência de elementos, o procedimento *achaMediana()* é utilizado. A idéia é que o ponto de partição da página distribua a área $r.w \times r.h$ disponível de forma balanceada entre as duas regiões filhas. Este procedimento é ilustrado no Algoritmo 1.

Quanto à decisão de qual eixo escolher para particionar uma região (horizontal ou vertical), o algoritmo opta pelo eixo mais longo, exceto quando não há colunas disponíveis para uma par-

Algoritmo 1 Procedimento para localizar o índice de partição de uma região.

Algoritmo *achaMediana* (i, j)
 $minerro \leftarrow \infty$
para $kn = i$ a $j - 1$ **faça**
 $err \leftarrow \max \left(\sum_{l=i}^{kn} A_l, \sum_{l=kn+1}^j A_l \right)$
se ($err < minerro$) **então**
 $minerro \leftarrow err$
 $minkn \leftarrow kn$
fim se
fim para
retorne $minkn$

tição vertical. Neste caso, o algoritmo particiona horizontalmente. A tendência neste processo de particionar o eixo mais longo é que a razão de aspecto de todas as regiões tendem a ser mais uniformes entre si, e se aproximar de uma forma quadrada, isto é, $AR = 1$.

No caso de uma partição horizontal, a altura da região acima do elemento k onde $i \leq k < j$, em uma região de altura H_r , é calculada como

$$H_{acima} = H_r \times \frac{\sum_{l=i}^k A_l}{\sum_{l=i}^j A_l}, \quad (4.11)$$

ou seja, é a altura proporcional às áreas na esquerda de k . Conseqüentemente, a área abaixo de H_{acima} é o complemento

$$H_{abaixo} = H_r - H_{acima}. \quad (4.12)$$

Já em uma partição vertical, escolhe-se o número de colunas mais próximo da mediana, calculada anteriormente. O Algoritmo 2 descreve este procedimento.

Algoritmo 2 Procedimento para selecionar o número de colunas para uma partição vertical.

Algoritmo *achaNumColunas* ($c, c_{esq}, c_{dir}, proporcao$)
 $minErro \leftarrow \infty$
 # *Encontra o corte vertical mais próximo da proporção calculada*
para $kc = c_{esq}$ a $c - c_{dir}$ **faça**
 $err \leftarrow |kc/c - proporcao|$
se ($err < minErro$) **então**
 $minErro \leftarrow err$
 $minkc \leftarrow kc$
fim se
fim para
retorne $minkc$

Visto que imagens possuem um número fixo de colunas, o algoritmo também leva em consideração um corte possivelmente longe da mediana, mas que atende o número de colunas desejado para uma imagem. Isto é feito selecionando-se os maiores números de colunas c_{esq}

e c_{dir} de uma imagem à esquerda da mediana $l \leq k$ e à direita da mediana $k < m \leq j$, e aceitando uma partição vertical somente quando o número de colunas c da região for maior ou igual à $c_{esq} + c_{dir}$, garantindo o número mínimo de colunas pedido pelas imagens no intervalo $[i \dots j]$.

O algoritmo principal de bissecção é descrito em detalhes no Algoritmo 3 e a Figura 9 ilustra o resultado do processo de bissecção dividido em uma série de passos. A chamada inicial para o algoritmo é *Bisseccao*(*raiz*, 1, *n*, *NC*, *W*, *H*, 0, 0). Vale observar que algumas das etapas do algoritmo ainda serão descritas no decorrer deste capítulo, como a medida de erro por exemplo (Subseção 4.5.1).

Algoritmo 3 O algoritmo de bissecção principal.

Pré-condição: Lista de elementos E_1, \dots, E_n normalizada com áreas A_1, \dots, A_n

Algoritmo *Bisseccao*(*ptr*, *i*, *j*, *c*, *w*, *h*, *x*, *y*)
 # armazena os atributos da região em *ptr*
ptr. {*i*, *j*, *c*, *w*, *h*, *x*, *y*} ← {*i*, *j*, *c*, *w*, *h*, *x*, *y*}
 # Se $i = j$, então *ptr* é uma região terminal.
se ($i = j$) **então**
 retorne erro(*i*, *c*, *w*, *h*)
fim se
minkn ← *achaMediana*(*i*, *j*)
proporcao ← $\sum_{l=i}^{minkn} A_l / \sum_{l=i}^j A_l$
c_{esq} ← $\max \{C_l \mid l \in [i; minkn], \text{imagem}(l)\}$
c_{dir} ← $\max \{C_l \mid l \in [minkn + 1; j], \text{imagem}(l)\}$
se $h > w$ ou $c < c_{esq} + c_{dir}$ **então**
 # Corte HORIZONTAL: Divide recursivamente as regiões acima e abaixo de h_{acima}
 h_{acima} ← $h \times proporcao$
 erroEsq ← *Bisseccao*(*ptr.esq*, *i*, *minkn*, *c*, *w*, *h_{acima}*, *x*, *y*)
 erroDir ← *Bisseccao*(*ptr.dir*, *minkn* + 1, *j*, *c*, *w*, $h - h_{acima}$, *x*, $y + h_{acima}$)
senão
 # Corte VERTICAL: Divide recursivamente as regiões à esquerda e direita de w_{esq}
 minkc ← *achaNumColunas*(*c*, *c_{esq}*, *c_{dir}*, *proporcao*)
 w_{esq} ← $w \times minkc / c$
 erroEsq ← *Bisseccao*(*ptr.esq*, *i*, *minkn*, *minkc*, *w_{esq}*, *h*, *x*, *y*)
 erroDir ← *Bisseccao*(*ptr.dir*, *minkn* + 1, *j*, $c - minkc$, $w - w_{esq}$, *h*, $x + w_{esq}$, *y*)
fim se
 # O erro desta região é a soma dos erros das regiões-filhas
retorne *erroEsq* + *erroDir*

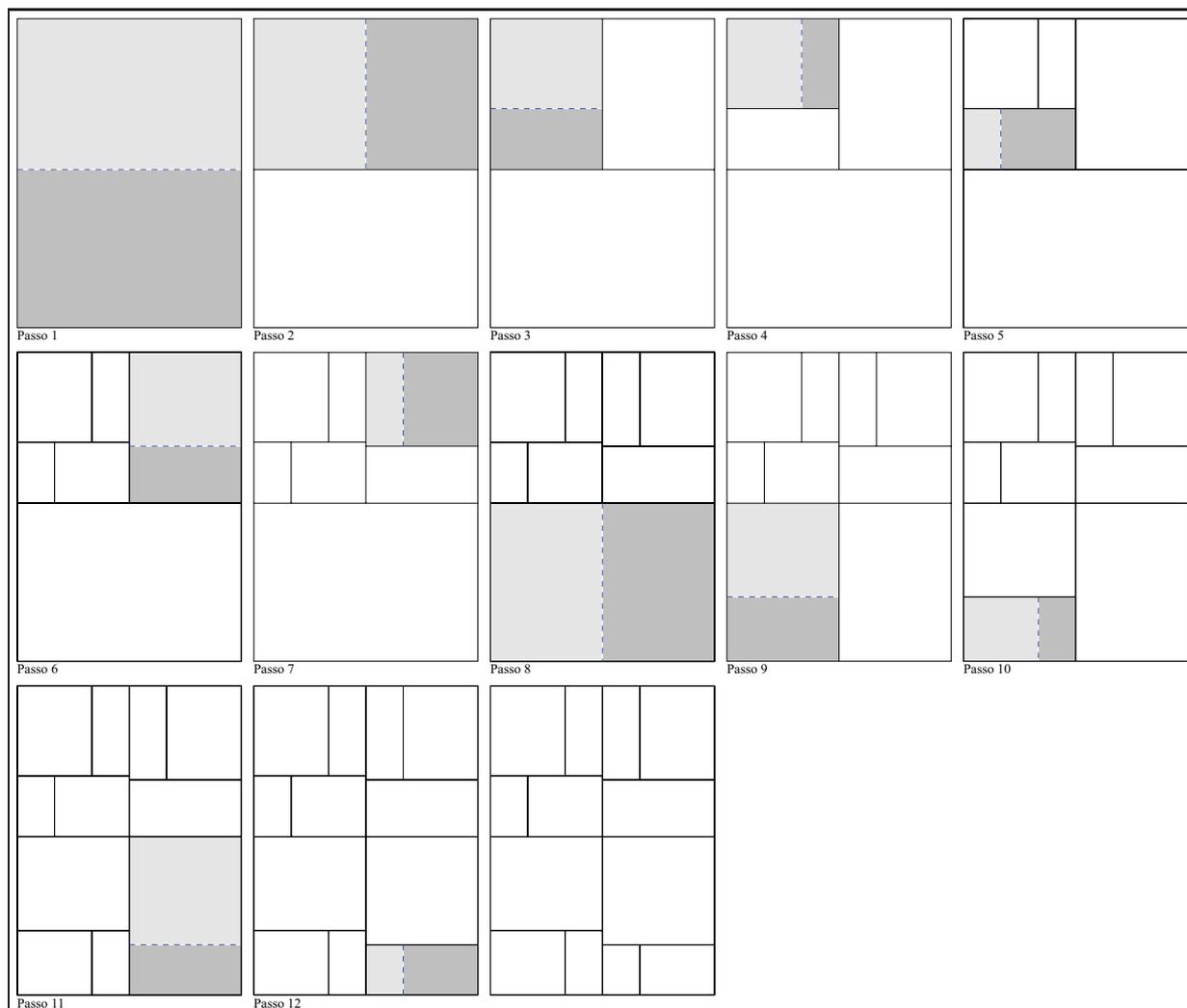


Figura 9: Um exemplo do algoritmo de bissecção em uma página de documento.

4.3.7 Renderização e geração do documento final

Após a execução do algoritmo de bissecção, a estrutura de regiões produzida pode ser processada e convertida em comandos para um formatador de texto externo, no caso deste trabalho, o \LaTeX [39].

Para produzir documentos com aparência similar a uma página de jornal, foram desenvolvidas algumas funcionalidades para o \LaTeX , tais como correções tipográficas e definições estilísticas para o modelo de um jornal (por exemplo, emprego de títulos, autores, números de página, etc.). As principais funcionalidades são descritas a seguir.

Colunas lógicas *versus* colunas visuais

Segundo Harrower [15], o impacto visual de um jornal é bastante influenciado pela divisão do texto em colunas. Como o algoritmo desenvolvido neste trabalho divide o documento em um

número fixo de colunas de mesma largura, um texto que ocupe muitas colunas apresentado na impressão final do documento pode ser difícil de ler ou ser aparentemente desagradável, devido ao excesso de regularidade na página. Por esta razão, uma das correções feitas na etapa de renderização neste trabalho é a redução “visual” do número de colunas de uma página, porém, o texto é apresentado com um número menor de colunas. Mais especificamente, o número de colunas visuais de um texto é a metade do número de colunas lógicas, quando um texto ocupa mais de duas colunas. A Figura 10 ilustra este efeito.



(a) Um documento produzido sem a redução visual de colunas.



(b) O mesmo documento, porém, com o emprego da redução visual de colunas.

Figura 10: Exemplo do emprego de colunas visuais em um documento com um número fixo de colunas lógicas.

Encaixe de regiões de texto

Na produção do documento final, as áreas calculadas para os textos podem ser maiores ou menores do que as áreas requisitadas por estes, fazendo com que os textos ocupem menos espaço na página, ou atravessando a região que o contém, por serem maiores. Sendo um algoritmo heurístico, o preenchimento exato de um texto em uma região não é garantido, apenas aproximado. Logo, com este sistema, é necessário que o formador de textos se encarregue de fazer

com que os textos ocupem suas regiões de forma precisa, a fim de garantir a homogeneidade da página.

Para resolver este problema, uma solução similar à proposta por Hurst et al. [42] foi adotada: assumindo uma região retangular fixa, o espaçamento entre linhas e tamanho de fonte do texto contido nesta região é alterado repetidamente de forma a aumentar ou diminuir a densidade de texto contida na região, até que o texto ocupe toda a região que o contém, sem exceder o seu limite. Esta solução foi desenvolvida com o \LaTeX , e é ilustrada na Figura 11.

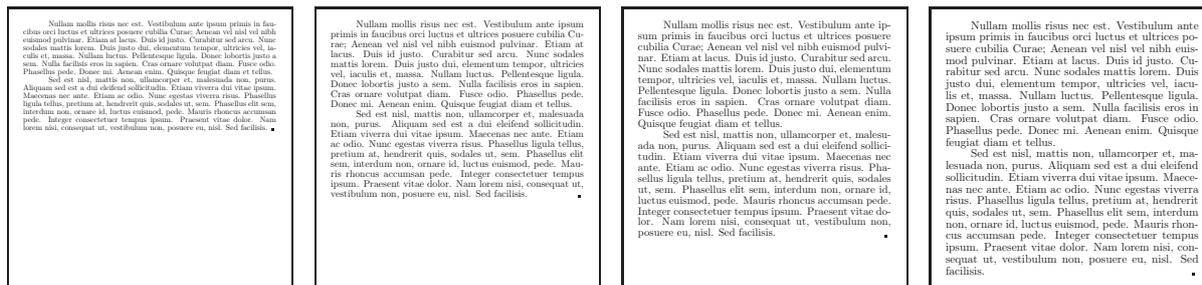


Figura 11: Pós-processamento nas regiões de texto para garantir a homogeneidade da página. Neste caso, o tamanho da fonte e espaçamento entre linhas é aumentado até que o texto ocupe toda a região.

4.3.8 Preservação da ordem de leitura

Para documentos provenientes de culturas ocidentais – que é o modelo seguido neste trabalho – a ordem de leitura em um documento é dada de cima para baixo, e da esquerda para a direita [16]. Porém, como discutido por Meunier [16], esta descrição descreve uma ordem *parcial* apenas, pois, para um dado elemento da página, o elemento lido anteriormente pode estar tanto acima quanto à esquerda deste. Apesar desta ambigüidade, um leitor humano normalmente é capaz de identificar a ordem correta de leitura, auxiliado pelo conteúdo do texto [16].

Seguindo esta definição, o algoritmo proposto neste trabalho garante sempre a preservação da ordem de leitura de entrada em um *layout* de documento, pois como uma determinada região é particionada em duas regiões em um ponto k da seqüência ordenada de elementos, os elementos $(E_i \dots E_k)$ estarão em uma região acima (caso tenha ocorrido uma bissecção horizontal), ou à esquerda (caso tenha ocorrido uma bissecção vertical) da seqüência $(E_{k+1} \dots E_j)$. Quando o algoritmo termina, produzindo uma região por elemento em uma página, é garantido que para cada elemento E_k seus predecessores estão acima ou à esquerda em relação à página do documento, o que obedece à definição de ordem de leitura utilizada neste trabalho.

A Figura 12(a) mostra um documento gerado com 15 elementos de texto e o mesmo documento é mostrado na Figura 12(b), porém, uma linha azul cruzando o documento indica a ordem de leitura, assim como a intensidade de cada região da página, ou seja, quanto mais es-

cura uma região, mais para o fim da seqüência estará o elemento contido nesta região. Note que esta ordem respeita a regra descrita nesta seção.

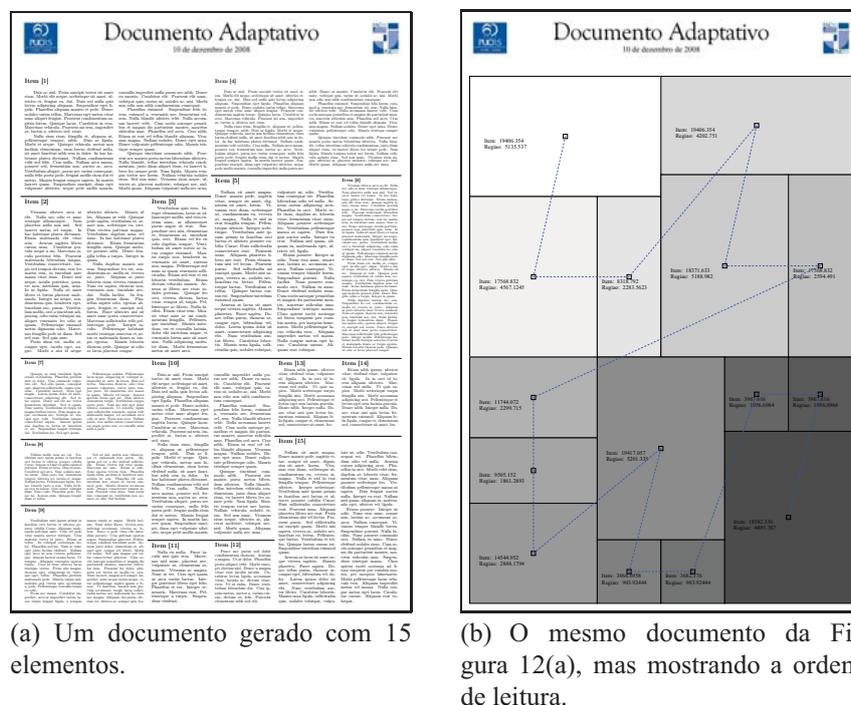


Figura 12: Exemplo de preservação da ordem de leitura

4.4 Paginação do documento

Para distribuir a seqüência de elementos em diversas páginas, foi adaptado o método descrito por Skiena [43], que é capaz de particionar uma seqüência de j elementos em k grupos, de forma que as seqüências resultantes dos grupos tenham as áreas mais similares quanto possível. A recorrência que descreve o algoritmo é dada como

$$M_{j,k} = \begin{cases} A_j & j = 0 \\ \sum_{l=0}^j A_l & k = 1 \\ \min_{0 \leq i \leq j-1} \max \left(M_{i,k-1}, \sum_{l=i+1}^j A_l \right) & \end{cases} \quad (4.13)$$

Esta recorrência (Equação (4.13)) produz divisões com áreas mais similares o possível, porque através da minimização da área do maior grupo em cada divisão as áreas dos grupos menores devem ser menores do que esta maior área, produzindo grupos com áreas o mais similares quanto possível.

Para fins de simplificação, a idéia proposta neste trabalho é a de tratar a paginação como um processo independente do algoritmo de *layout*, ou seja, dado um número k desejado de páginas,

a seqüência de elementos é dividida em k seqüências menores, através do método descrito na Equação (4.13). As seqüências resultantes possuem suas áreas totais o mais similares entre si quanto possível, e para cada uma é executado o algoritmo de bissecção de forma independente. Esta estratégia, apesar de não-ótima, garante a divisão homogênea de elementos através das páginas resultantes.

Nas Figuras 13, 14, 15 e 16 é apresentado um exemplo, em que uma página com uma seqüência de 10 elementos é dividida em 1, 2, 3, e 4 páginas, respectivamente. Em cada caso, é mostrada a seqüência de elementos e a área total em cada grupo (normalizadas e multiplicadas por 1000, para facilitar a visualização), nas Tabelas 1, 2, 3 e 4.

Tabela 1: Áreas dos elementos em uma página.

Núm. página	Áreas dos elementos						Área total (Σ)				
1	181	152	30	41	82	38	95	118	118	145	1000
Total (Σ):											1000



Figura 13: Um documento de uma página contendo 10 elementos.

Tabela 2: Áreas divididas em duas páginas.

Núm. página	Áreas dos elementos						Área total (Σ)
1	181	152	30	41	82		486
2		38	95	118	118	145	514
Total (Σ):							1000



Figura 14: Um documento gerado com 10 elementos dividido em duas páginas.

Tabela 3: Áreas divididas em três páginas.

Núm. página	Áreas dos elementos			Área total (Σ)
1	181	152	30	363
2	41	82	38 95 118	374
3	118	145		263
Total (Σ):				1000



Figura 15: Um documento gerado com 10 elementos dividido em três páginas.

Tabela 4: Áreas divididas em quatro páginas.

Núm. página	Áreas dos elementos			Área total (Σ)
1	181			181
2	152	30	41 82	305
3	38	95	118	251
4	118	145		263
Total (Σ):				1000



Figura 16: Um documento gerado com 10 elementos dividido em quatro páginas.

4.5 Melhorando os resultados gerados

Como o algoritmo apresentado testa apenas uma forma de realizar bissecções na página, o *layout* gerado pode não ser esteticamente agradável. Para determinar a qualidade do *layout* gerado para um documento, introduz-se uma função de erro, descrita a seguir na Subseção 4.5.1. Através da variação de alguns parâmetros na entrada do documento, é possível executar o algoritmo de bissecção diversas vezes, e escolher os parâmetros que foram responsáveis pelo menor erro gerado na página. A seguir é apresentada a medida de erro para os *layouts* gerados pelo algoritmo.

4.5.1 Medida de erro da página

O método ilustrado no Algoritmo 3 utiliza uma função de erro como retorno do procedimento de particionamento recursivo. Este valor pode ser utilizado por novas etapas que coordenem a execução do algoritmo múltiplas vezes, a fim de buscar soluções melhores.

Para medir o erro de um *layout* gerado para um documento com n elementos, pode-se realizar uma soma dos erros gerados em cada elemento, como apresentado no Algoritmo 3. O erro de um texto com área normalizada A_i que foi alocado para uma região R com área $A_R = R.w \times R.h$ e razão de aspecto $AR_R = R.w/R.h$ é definido como

$$\max \left(\frac{A_R/(W \times H)}{A_i}, \frac{A_i}{A_R/(W \times H)} \right), \quad (4.14)$$

ou seja, o tamanho da proporção entre as áreas A_i e A_R , relativas à área da página $W \times H$ (Seção 4.2).

Para medir o erro gerado em imagens, que possuem geometria fixa, pode-se usar o tamanho da proporção entre as razões de aspecto da imagem AR_i e da região, ou seja

$$\max \left(\frac{AR_R}{AR_i}, \frac{AR_i}{AR_R} \right). \quad (4.15)$$

Note que as Equações (4.14) e (4.15) medem a proporção de diferença entre as áreas/aspectos requisitados por um elemento e as áreas/aspectos fornecidos no *layout* final para os mesmos, sendo 1 o valor ideal (ausência de erro). Para uma medida mais precisa, observe que o erro deve aumentar à medida que a área do elemento aumenta na página, ou seja, o erro deve ser proporcional à área da região calculada. Ainda, a ausência de erro deve ser medida como 0, por isso subtrai-se 1 do valor medido. Portanto, a medida de erro *err* deve ser mapeada como

$$(err - 1) \times w \times h, \quad (4.16)$$

onde w e h são a largura e a altura da região, respectivamente (Seção 4.3.5).

O Algoritmo 4 descreve a função de erro utilizada pelo Algoritmo 3.

Algoritmo 4 Função para medida de erro de uma página para o algoritmo.

Algoritmo $erro(i, c, w, h)$
 $area \leftarrow w \times h$
se $imagem(i)$ **então**
 $AR_{regiao} \leftarrow w/h$
 $err \leftarrow \max(AR_{regiao}/AR_i, AR_i/AR_{regiao})$
senão
 $err \leftarrow area/W \times H$
 $err \leftarrow \max(err/A_i, A_i/err)$
fim se
retorne $(err - 1) \times area$

4.5.2 Redimensionando imagens

Se for permitida a variação do número de colunas C_j informado como entrada para a imagem j (Subseção 4.3.3), é possível testar diversos *layouts* diferentes, um para cada largura diferente para cada imagem. Para explorar todas as possibilidades, é executado um algoritmo de *backtracking* [43] que seleciona a combinação (C_1, C_2, \dots, C_m) para as imagens $1 \dots m$ que provê o menor erro de *layout* gerado (descrito na Subseção 4.5.1).

4.5.3 Mantendo elementos agrupados

Como requisito em documentos como jornais, descritos na Seção 4.2, é necessário que elementos relacionados entre si apareçam próximos uns dos outros na página final. A simples preservação da ordem de leitura não é suficiente, pois se uma bissecção em um ponto k no intervalo $[i \dots j]$ divide a seqüência em $[i \dots k]$ e $[k + 1 \dots j]$ (Seção 4.3.6), não é garantido que os elementos k e $k + 1$ sairão próximos na página, graças ao processo de bissecção. A Figura 17 ilustra este problema.

No caso da Figura 17, a imagem no canto esquerdo superior da página deveria aparecer visualmente alinhada à notícia acima (“*Imagem da boa mãe...*”), porém, o *layout* produzido causa uma descontinuidade visual entre os dois elementos, fazendo com que o leitor acredite que não há conexão lógica entre os mesmos.

Para resolver este problema, é possível transformar a seqüência de elementos da entrada do algoritmo em uma estrutura hierárquica, onde nós não-terminais representam grupos em que os componentes de um determinado grupo são os nós filhos deste. Um nó terminal representa

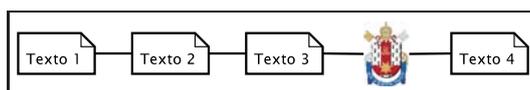


(a) Representação da lista de elementos de um documento.

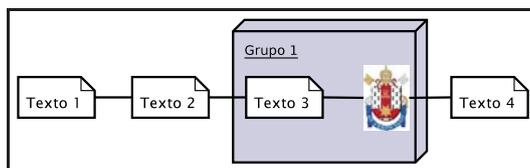
(b) Um documento produzido automaticamente, sem agrupamento lógico entre elementos.

Figura 17: Um documento sem grupos ilustrando a desconexão visual entre o texto e a imagem no canto superior da página.

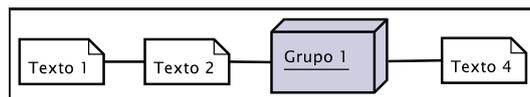
um elemento do documento (texto ou imagem). A Figura 18 ilustra o processo de agrupamento sobre uma lista simples de elementos.



(a) Lista linear de elementos.



(b) Adição de um grupo para os elementos *Texto2* e *LogoPUC*.



(c) Representação em lista para o algoritmo de bissecção. Os grupos são processados recursivamente após o *layout* ter sido gerado para o nível superior.

Figura 18: Processo de agrupamento de elementos em uma lista.

Para tornar essa hierarquização possível, é necessário adaptar o algoritmo de bissecção des-

crito na Seção 4.3.6 para que este processe elementos agrupados. A idéia é que cada grupo na hierarquia seja considerado um elemento indivisível e o algoritmo seja aplicado recursivamente para cada grupo toda vez que completar a execução para uma lista de elementos.

Resolvendo o problema apresentado na Figura 17 com agrupamento produz os resultados desejados, como pode ser visto na Figura 19.



(a) Representando grupos através da hierarquização da lista de elementos.

(b) Um documento produzido automaticamente, com agrupamento lógico entre elementos.

Figura 19: O mesmo documento da Figura 17, porém com agrupamento lógico entre elementos.

4.6 Objetivos atingidos com o método proposto

O método desenvolvido, graças às restrições quanto ao formato do documento é razoavelmente simples, mas ainda é possível identificar algumas contribuições para problemas gerais na área de VDP (Capítulo 2), que puderam ser realizadas com o método descrito neste capítulo:

- Obtenção de um algoritmo determinístico e com boa performance (através do algoritmo de divisão e conquista obtém-se o limite superior $O(n \times \log_2(n))$ no caso médio, não explicado neste trabalho) para realizar a alocação de elementos em uma página de documento;
- O algoritmo é capaz de preservar critérios estéticos importantes, como alinhamento (vertical) dos elementos e uniformidade (ver Capítulo 2);

- A ordem de leitura da seqüência de entrada é preservada (Subseção 4.3.8);
- Elementos podem ser agrupados, como descrito na Subseção 4.5.3;
- Através da especificação do número de páginas desejado, é possível dividir um documento em uma ou mais páginas, como descrito na Subseção 4.4;
- O método descrito neste trabalho pode ser estendido para soluções mais sofisticadas, que explorem um espaço de soluções maior, guiados por heurísticas como a medida de erro/custo e o método para redimensionar imagens descritos na Seção 4.5;
- É possível misturar elementos com aspecto fixo (imagens) com texto, que possui forma livre (podendo possuir qualquer aspecto para preencher uma área da página), ao contrário da maioria dos trabalhos existentes (ver Capítulo 3) até o momento da publicação deste trabalho;

No Capítulo 5 a seguir serão apresentados alguns resultados deste trabalho, que foram produzidos através de uma avaliação empírica sobre a qualidade e performance dos resultados gerados.

5 Avaliação empírica e resultados

Um sistema capaz de produzir diagramações automaticamente é passível de produção de maus resultados, devido à dificuldade de se quantificar ou abstrair o que constitui uma “boa” diagramação de forma computacional [3]. Ainda, como discutido na Seção 2.3, o processo de impressão VDP exige que as soluções de diagramação automática executem em um curto período de tempo, a fim de não congestionar o fluxo contínuo de impressão de documentos. A fim de avaliar os resultados e a performance do algoritmo proposto neste trabalho (Capítulo 4), foram definidos dois experimentos para avaliar a qualidade e, em menor extensão, a performance do algoritmo descrito neste trabalho. Desta forma, a avaliação será dividida em duas etapas:

- **Avaliação qualitativa:** foi definido um experimento onde uma amostra de pessoas especializadas em *design* gráfico deveria avaliar uma série de documentos produzidos automaticamente pelo algoritmo proposto neste trabalho;
- **Testes de performance:** o algoritmo desenvolvido foi executado para um determinado número de documentos gerados de forma aleatória como descritos na Seção 5.2, porém contendo uma quantidade incremental de elementos, colunas e figuras a fim de avaliar o tempo de execução do algoritmo para documentos mais sofisticados e identificar possíveis problemas de performance no algoritmo ou nas tarefas complementares (Capítulo 4);

Este capítulo portanto descreve os experimentos realizados e os resultados obtidos desta avaliação. A Seção 5.1 descreve o experimento para a avaliação qualitativa, em seguida apresentando os resultados obtidos e a interpretação dos mesmos. A Seção 5.2, de maneira similar, apresenta os testes de performance realizados sobre o algoritmo.

5.1 Avaliação qualitativa

5.1.1 Descrição do experimento realizado

Como apresentado no Apêndice A, o teste é composto por 20 casos de teste. Cada caso de teste é um documento produzido automaticamente pelo algoritmo proposto neste trabalho, simulando capas de jornal (conforme descrito na Seção 4.2) e uma avaliação por nota (descrita no decorrer desta seção) em um único critério em cada documento, que é o seu aspecto geral em termos de qualidade (diagramação, distribuição de áreas na página, etc.).

Os textos e imagens para cada caso foram obtidos de uma base de dados contendo 5841 elementos igualmente distribuídos no total, constituindo 4866 textos e 975 imagens.¹ Os textos contidos nesta base possuem aproximadamente entre 10 e 3000 caracteres de tamanho, distribuídos de maneira aproximadamente uniforme neste intervalo. A geração de casos de teste foi feita a partir de uma seleção aleatória de elementos desta base, distribuídos de maneira uniforme e obedecendo aos seguintes critérios, considerados valores típicos para publicações de jornais tradicionais [15]:

- O número de elementos varia de 5 a 12;
- O número de colunas varia entre 5 e 7;
- O percentual de elementos que são figuras é de 25%, sendo que pelo menos dois casos devem conter apenas texto.

Os casos de teste gerados podem ser visualizados no Apêndice A e são descritos a seguir, na Tabela 5.

Tabela 5: Casos de teste para comparação por medidas estéticas.

Caso	Núm. elementos	Núm. colunas	% São imagens
1	12	7	Sem imagens
2	9	7	25%
3	12	6	Sem imagens
4	10	6	25%
5	9	6	25%
6	8	6	25%
7	7	6	25%
8	6	6	25%
9	5	6	25%
10	5	5	25%

Reordenação de elementos dos casos de teste

A fim de facilitar a identificação posterior de problemas na diagramação, foi levantada uma hipótese quanto à ordem dos elementos de uma página. Como a ordem de entrada dos elementos se mantém fixa durante a execução do algoritmo, sabe-se pela descrição na Seção 4.3.6 que uma ordem de entrada diferente produzirá uma diagramação diferente para o mesmo documento, com uma qualidade possivelmente melhor ou pior do que a original. Portanto, para cada um dos 10 casos de teste da Tabela 5 foi gerada uma variação a partir da mudança de ordem de alguns elementos da seqüência do documento, consideradas aqui como melhorias sobre o caso de teste original correspondente. Desta forma, a avaliação é constituída por 20 casos de teste

¹Os textos e imagens foram obtidos na *internet*, através dos serviços RSS [44] *Terra* [45] e *Folha de São Paulo* [46].

no total. É importante ressaltar que, para evitar respostas tendenciosas nas avaliações, a informação de que os últimos 10 casos de teste eram versões reordenadas dos 10 primeiros casos foi omitida aos voluntários. Além disto, há dois casos de teste (Caso 1 e Caso 10) em que a ordem não foi modificada, sendo estes simplesmente replicados, a fim de identificar situações em que os voluntários possam ter preenchido respostas influenciadas por documentos avaliados anteriormente.

Os voluntários escolhidos como amostra foram alunos da FAMECOS da Pontifícia Universidade Católica do Rio Grande do Sul. Inicialmente, a amostra era composta por cerca de 42 alunos com experiência em pelo menos uma disciplina de projeto gráfico de documentos. Porém, apenas 13 voluntários responderam ao teste, prejudicando a avaliação de certa forma, como será discutido ao final desta seção.

O teste aplicado para cada voluntário é dividido nas seguintes etapas:

1. Um texto explicativo, ilustrado na Figura 42 instrui o voluntário para o teste que será realizado, que envolve a classificação quanto à qualidade de uma série de documentos produzidos automaticamente;
2. Antes de avaliar as páginas, o voluntário deve responder a um questionário pré-teste (ilustrado na Figura 43), contendo perguntas relacionadas a experiência do mesmo com o projeto gráfico de documentos e outras informações demográficas, como idade e sexo, a fim de facilitar a classificação posterior da amostra;
3. Após o questionário, são apresentados os vinte casos de teste ao voluntário, que avaliará cada um e preencherá uma grade (Figura 44) com uma nota na escala *Likert* [47] de 1 a 5 (“muito ruim” a “muito bom”), com o valor que este achar adequado à qualidade da diagramação do caso de teste respectivo.

Com o experimento descrito, a seguir será apresentado o resumo dos dados da amostra escolhida para a avaliação qualitativa, para então apresentar os resultados obtidos da avaliação.

5.1.2 Descrição da amostra de voluntários

A Tabela 6 apresenta os dados da amostrada coletados do questionário pré-teste. Note que, devido ao tamanho pequeno da amostra, decidiu-se por não resumir estes dados, sendo estes apresentados de forma integral.

Tabela 6: Dados da amostra coletados no questionário pré-teste (Figura 43).

Idade	Sexo	Frequência de leitura de jornais (por semana)	Publicações diferentes lidas	Semestres na faculdade	Disciplinas de projeto gráfico cursadas	Experiência profissional
19	M	mais de 3 vezes	mais de 3	4	2	Até 1 ano
21	F	mais de 3 vezes	1	8	4	Nenhuma
21	F	mais de 3 vezes	2 a 3	6	3	Nenhuma
21	M	1 vez	2 a 3	7	3	Até 1 ano
21	M	3 vezes	1	5	2	Nenhuma
21	M	3 vezes	1	5	3	Nenhuma
22	F	mais de 3 vezes	2 a 3	6	3	Até 1 ano
24	F	mais de 3 vezes	2 a 3	10 ou mais	3	Nenhuma
24	F	mais de 3 vezes	2 a 3	8	30	De 1 a 2 anos
25	M	1 vez	1	10 ou mais	6	Nenhuma
26	F	1 vez	1	6	1	Nenhuma
28	M	3 vezes	2 a 3	3	1	Mais de 2 anos
35	M	mais de 3 vezes	mais de 3	10 ou mais	30	Mais de 2 anos

5.1.3 Descrição dos resultados

Os resultados da avaliação por parte dos voluntários estão resumidos no gráfico da Figura 20 e a Tabela 7 apresenta as médias e os valores mais frequentes (modas) das respostas dadas para cada caso de teste.

Tabela 7: Médias e modas das respostas dos voluntários para cada caso de teste, incluindo a respectiva versão reordenada do documento.

Caso de teste	Média	Média c/ reordenação	Moda	Moda c/ reordenação
1	Ruim	Razoável	Ruim	Ruim
2	Ruim	Razoável	Ruim	Ruim
3	Ruim	Ruim	Razoável	Ruim
4	Ruim	Razoável	Ruim	Razoável
5	Razoável	Razoável	Ruim	Ruim
6	Razoável	Razoável	Ruim	Razoável
7	Razoável	Razoável	Ruim	Razoável
8	Ruim	Razoável	Ruim	Razoável
9	Ruim	Razoável	Ruim	Razoável
10	Razoável	Razoável	Bom	Razoável

Na Figura 20 há 20 gráficos de barra. Os gráficos na parte positiva do eixo representam os 10 primeiros casos de teste gerados aleatoriamente, e os da parte negativa representam os últimos 10 casos de teste, que são versões cuja ordem dos elementos foi modificada artificialmente na tentativa de obter um melhor resultado qualitativo. A seguir será realizada uma interpretação dos resultados obtidos com esta avaliação.

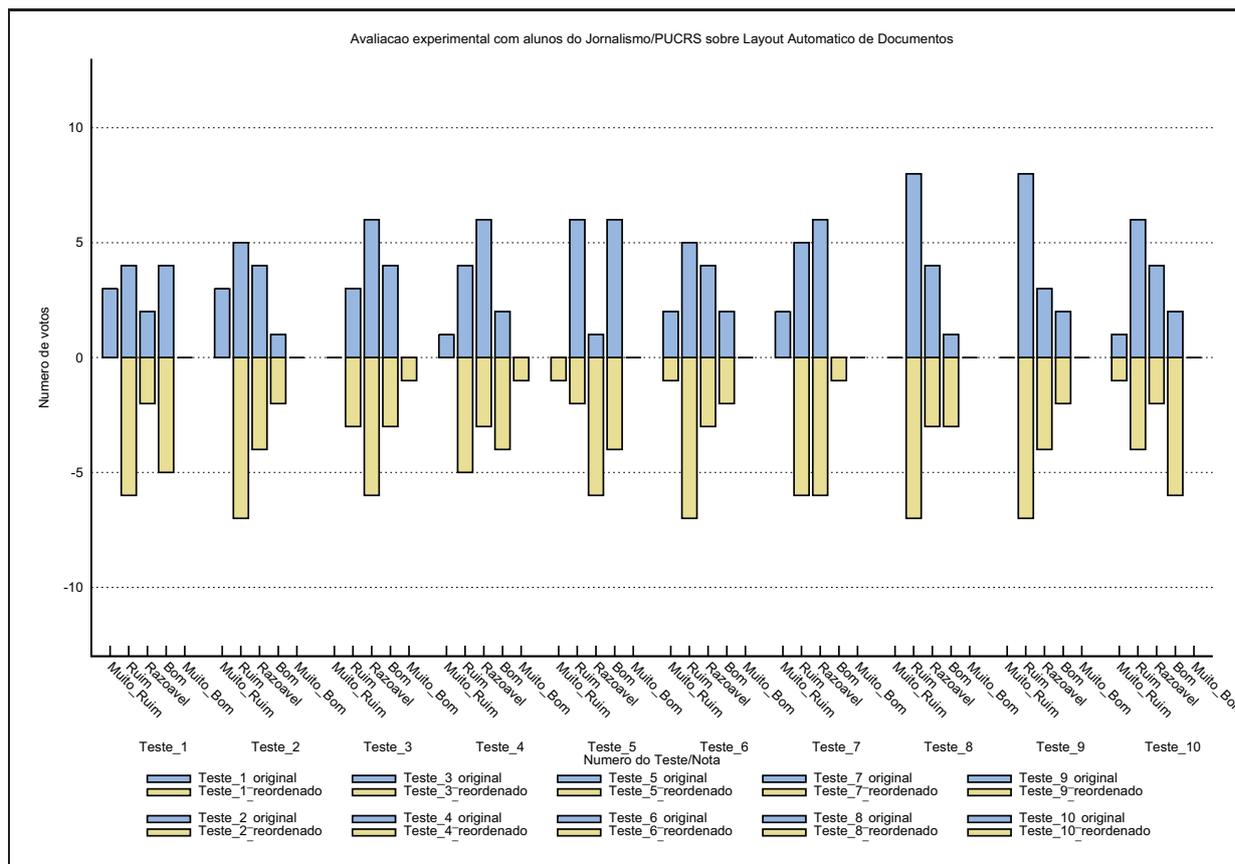


Figura 20: Resultados gerados pela avaliação experimental.

5.1.4 Interpretação dos resultados

Devido ao baixo número de pessoas que fizeram o teste, não foi possível obter resultados conclusivos, mas é possível identificar pela Tabela 7 que a simples mudança de ordem nos elementos feita em cada caso de teste colaborou para a melhoria dos resultados obtidos. Isto indica que a ordem de entrada dos elementos de um documento afeta a qualidade do resultado gerado. Isto pode sugerir que, se a ordem de leitura não for um requisito obrigatório, a construção de um mecanismo que calcule novas ordens para os elementos sobre o algoritmo original seria uma melhoria possível para o algoritmo.

Analisando os comentários e observações preenchidos pelos voluntários, percebeu-se que estes foram muito críticos em relação a critérios além do escopo da diagramação do documento, como tamanhos de fontes, parágrafos justificados, etc. que estão além do controle do algoritmo, pois são parâmetros controlados pelo formatador de texto utilizado (Seção 4.3.7). Uma possível causa deste desvio de foco nas críticas foi o escopo demasiado geral da pergunta proposta aos voluntários (ver Figura 42), sem questionar a diagramação de cada página de forma explícita. Estes fatores certamente colaboraram para os resultados em sua maioria ruins. Apesar disto, as observações feitas pelos voluntários quanto à diagramação do documento foram positivas em geral, contribuindo para o número de respostas positivas nos resultados.

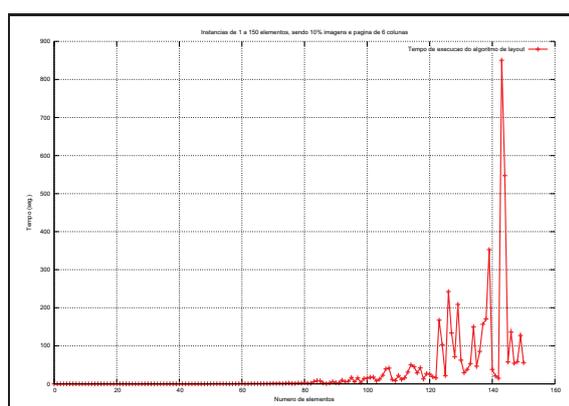
Apesar de a qualidade dos documentos gerados pelo algoritmo proposto neste trabalho não ser garantida todo o tempo, é possível aplicá-lo para a construção de documentos de *vida curta*, tais como jornais diários e *newsletters*, onde a qualidade superior de um projeto gráfico manual não é tão importante quanto publicar o conteúdo no tempo devido. Isto é, os resultados produzidos pelo algoritmo desenvolvido neste trabalho são passíveis de problemas estéticos, tais como muita variação no tamanho de fontes entre dois elementos distintos. Porém, com um conteúdo selecionado apropriadamente este tipo de problema não acontece com frequência e o *layout* gerado pode ser considerado “bom o bastante”, se contrastado com o baixo custo de produção, comparado ao custo de um documento produzido manualmente.

5.2 Testes de performance

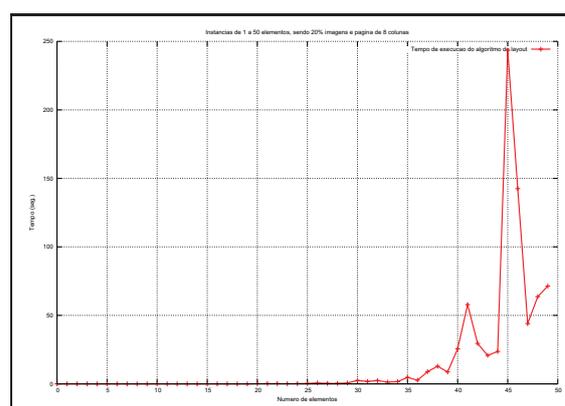
Neste experimento o algoritmo desenvolvido neste trabalho será executado para um número de documentos gerados de forma aleatória, em duas breves etapas:

1. Instâncias de entradas para os algoritmos serão geradas com a mesma base descrita na Seção 5.1, porém com um número incremental de elementos variando entre 1 e 150 elementos em um documento com 6 colunas, sendo 10% dos elementos imagens, em média;
2. 50 casos são gerados, porém, com cada instância de documento contendo 8 colunas e sendo 20% dos elementos imagens, em média.

A Figura 21 apresenta os resultados obtidos com o tempo de execução, em segundos.



(a) Primeiro teste de performance.



(b) Segundo teste de performance.

Figura 21: Gráficos para a avaliação de performance do algoritmo de diagramação automática.

Analisando os gráficos na Figura 21, observa-se a manutenção de um tempo de execução curto (abaixo de um minuto) nas instâncias geradas até 120 elementos no primeiro teste e até 45 elementos no segundo teste. Verificando os algoritmos implementados, é possível identificar

que o aumento repentino e as flutuações observadas ocorrem devido ao algoritmo de redimensionamento de imagens (Seção 4.5), que possui um desempenho baixo quando o número de imagens e colunas é grande. Porém, para casos mais realistas, com um número de até 20 elementos, o algoritmo de diagramação automática demonstra ser razoavelmente eficiente.

O teste realizado não leva em consideração outras fases um ambiente de impressão de VDP, como ilustrado na Figura 1 (Capítulo 2), que visa aumentar a quantidade de páginas produzidas em uma quantidade de tempo determinada. Portanto, uma avaliação mais realista ainda é necessária, visto que o teste realizado aqui fornece apenas um indicador empírico do desempenho do algoritmo, independente do processo de impressão.

6 Considerações finais e trabalhos futuros

Foi proposto neste trabalho um novo algoritmo para diagramação automática de documentos adaptativos. O algoritmo segue o paradigma de divisão e conquista, realizando bissecções de forma recursiva em uma página. O ponto de bissecção e sua direção (horizontal ou vertical) são decididos através de heurísticas, que tentam distribuir as áreas de forma equilibrada pela página, assim como os aspectos de cada região. Além disso, diversas etapas adicionais e complementares ao algoritmo foram descritas nas Seções 4.5 e 4.4, como a paginação de um documento, o redimensionamento automático de imagens e agrupamento de elementos inter-relacionados, mostrando que o método proposto pode ser estendido para a produção de documentos mais sofisticados.

Com o método aqui desenvolvido foi feita uma tentativa de resolver problemas em métodos já existentes que, entre outros, são: a alocação de elementos considerando a ordem em que eles serão lidos no documento final, a preservação de critérios estéticos importantes para a qualidade visual do documento, e o desempenho (ex. aumentar a quantidade de páginas produzidas por minuto) exigido em aplicações de VDP (Capítulo 2). Apesar de o *layout* dos documentos gerados pelo algoritmo aqui proposto possuir uma forma simplificada (sendo dividido apenas em regiões retangulares, colunas e cobrindo a página inteira), consideramos os resultados como sendo bastante satisfatórios e encorajadores para um algoritmo determinístico – e baseado em heurísticas – e com performance (i. e., redução do tempo de execução) superior.

Como o método desenvolvido aqui é um algoritmo guloso [13], este sofre das mesmas limitações que outros algoritmos semelhantes, que é a produção de resultados não-ótimos devido ao conjunto pobre de escolhas feitas pelo algoritmo. Porém, se o alto desempenho não é um requisito em algum ambiente de aplicação deste algoritmo, uma melhoria que poderia trazer resultados significativamente melhores, seria tornar o algoritmo iterativo, realimentando a entrada com novas áreas toda vez que um *layout* fosse gerado, utilizando para isso a função de erro descrita na Seção 4.5.1. Porém, esta solução não foi experimentada, devido a preocupações quanto ao desempenho do algoritmo, e poderá ser investigada em trabalhos futuros.

Quanto à validação deste trabalho, visto que os resultados apresentados no Capítulo 5 foram prejudicados devido ao pequeno tamanho da amostra que participou da avaliação e às respostas influenciadas por fatores de formatação de texto nos casos de teste, torna-se necessária uma reavaliação qualitativa deste trabalho no futuro, considerando estratégias diferentes para selecionar e questionar as amostras. Uma avaliação quantitativa mais aprofundada também seria mais interessante, se aplicada em um ambiente de produção em massa de documentos, com impressoras de grande porte encarregadas da renderização do documento, visto que o maior gargalo

da solução aqui proposta seria durante a fase de renderização e formatação de texto, visto que foi observado um bom desempenho do algoritmo para casos mais realistas (Seção 5.2).

Outra possibilidade de validação seria utilizar os modelos descritos por Harrington et al. [3] e Cabeda e de Oliveira [10]. Neste trabalho isto não foi feito, devido às diferenças estéticas entre os documentos aqui produzidos e os documentos gerados em outros trabalhos. Ainda, *design* e critérios estéticos são passíveis de interpretação subjetiva, o que impediria conclusões definitivas quanto aos resultados deste trabalho.

Entre outras extensões possíveis para trabalhos futuros, pode-se citar:

- Realizar redimensionamento ou recorte automático nas imagens do documento [48], a fim de permitir a variação no aspecto destas, o que anula o erro de posicionamento de uma imagem em uma região por diferenças de aspecto;
- Estender o algoritmo para um método de otimização combinatória, a fim de garantir a melhor alocação de elementos no documento. Porém, deve ser levada em consideração a performance de tal método;
- Produzir interfaces para autoria de documentos adaptativos, em que o algoritmo de adaptação fornece visualizações em tempo real do documento para o usuário;
- Fazer reordenação iterativa de elementos, a fim de melhorar os resultados, como observado no Capítulo 5;
- Propor um método mais eficiente – possivelmente heurístico – para a reordenação de imagens, que atualmente é baseado em uma busca exaustiva por todas as maneiras possíveis de se definir o tamanho das imagens de um documento (Seção 4.5);
- Produzir variações do mesmo algoritmo, desconsiderando limitações como divisão em colunas, a fim de melhorar a alocação de elementos na página;
- Utilizar a geometria das regiões gerada pelo algoritmo de bissecção como *templates* iniciais para um método não-determinístico, como os citados neste trabalho [7–9], afim de melhorar o desempenho dos mesmos, com regiões pré-calculadas mais próximas de uma solução ótima.

Referências

- [1] BAGLEY, S. R.; BRAILSFORD, D. F.; OLLIS, J. A. Extracting reusable document components for variable data printing. In: *DocEng '07: Proceedings of the 2007 ACM symposium on Document engineering*. New York, NY, USA: ACM, 2007. p. 44–52.
- [2] SCHRIER, E. et al. Adaptive layout for dynamically aggregated documents. In: *IUI '08: Proceedings of the 13th international conference on Intelligent user interfaces*. New York, NY, USA: ACM, 2008. p. 99–108.
- [3] HARRINGTON, S. J. et al. Aesthetic measures for automated document layout. In: *DocEng '04: Proceedings of the 2004 ACM symposium on Document engineering*. New York, NY, USA: ACM Press, 2004. p. 109–111.
- [4] WILLIAMS, R. *The non-designer's design book: design and typographic principles for the visual novice*. Berkeley, CA, USA: Peachpit Press, 1994.
- [5] ATKINS, B. Adaptive photo collection page layout. In: *Image Processing, 2004. ICIP '04. 2004 International Conference on*. Washington, DC, USA: IEEE Computer Society, 2004. v. 5, p. 2897–2900 Vol. 5.
- [6] LIN, X. Active document layout synthesis. *Proceedings. Eighth International Conference on Document Analysis and Recognition, 2005.*, p. 86–90 Vol. 1, 29 Aug.–1 Sept. 2005.
- [7] GOLDENBERG, E. *Automatic layout of variable-content print data*. Dissertação (Mestrado) — School of Cognitive & Computing Sciences, University of Sussex, Brighton, UK, 2002. Disponível em: <<http://www.hpl.hp.com/techreports/2002/HPL-2002-286.html>>. Acesso em: 17 nov. 2009.
- [8] GEIGEL, J.; LOUI, A. *Automatic Page Layout Using Genetic Algorithms for Electronic Albuming*. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.18.7009>>. Acesso em: 17 nov. 2009.
- [9] PURVIS, L. et al. Creating personalized documents: an optimization approach. In: *DocEng '03: Proceedings of the 2003 ACM symposium on Document engineering*. New York, NY, USA: ACM, 2003. p. 68–77.
- [10] FARIA, A. C.; OLIVEIRA, J. ao B. S. de. Measuring aesthetic distance between document templates and instances. In: *DocEng '06: Proceedings of the 2006 ACM symposium on Document engineering*. New York, NY, USA: ACM Press, 2006. p. 13–21.
- [11] GOLDBERG, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1989.
- [12] JACOBS, C. et al. Adaptive grid-based document layout. In: *SIGGRAPH '03: ACM SIGGRAPH 2003*. New York, NY, USA: ACM Press, 2003. p. 838–847.

- [13] CORMEN, T. H. et al. *Introduction to Algorithms, Second Edition*. Cambridge, MA: The MIT Press, 2001. Paperback.
- [14] OLIVEIRA, J. B. S. de. Two algorithms for automatic document page layout. In: *DocEng '08: Proceeding of the eighth ACM symposium on Document engineering*. New York, NY, USA: ACM, 2008. p. 141–149.
- [15] HARROWER, T. *The Newspaper Designer's Handbook*. Boston, MA, USA: McGraw-Hill, 1992. 180 p.
- [16] MEUNIER, J.-L. Optimized xy-cut for determining a page reading order. In: *ICDAR '05: Proceedings of the Eighth International Conference on Document Analysis and Recognition*. Washington, DC, USA: IEEE Computer Society, 2005. p. 347–351.
- [17] BALDWIN, J.; ROWSON, J. A.; COADY, Y. Printmonkey: giving users a grip on printing the web. In: *DocEng '08: Proceeding of the eighth ACM symposium on Document engineering*. New York, NY, USA: ACM, 2008. p. 230–239.
- [18] PODI, *Personalized Print Markup Language (PPML) 2.2*. 2008. Disponível em: <<http://www.podi.org/>>. Acesso em: 10 dez. 2008.
- [19] GIANNETTI, F. A multi-format variable data template wrapper extending podis ppml-t standard. In: *DocEng '07: Proceedings of the 2007 ACM symposium on Document engineering*. New York, NY, USA: ACM, 2007. p. 37–43.
- [20] GIANNETTI, F. et al. High performance xsl-fo rendering for variable data printing. In: *SAC '06: Proceedings of the 2006 ACM symposium on Applied computing*. New York, NY, USA: ACM, 2006. p. 811–817.
- [21] PELICAN Digital – Variable Information Printing. 2008. Disponível em: <<http://www.pelicanpress.co.uk/digital/vip.html>>. Acesso em: 10 dez. 2008.
- [22] MENEGUZZI, F. R. et al. Strategies for document optimization in digital publishing. In: *DocEng '04: Proceedings of the 2004 ACM symposium on Document engineering*. New York, NY, USA: ACM, 2004. p. 163–170.
- [23] HP Indigo Digital Presses – HP – Graphic Arts. 2008. Disponível em: <<http://www.hp.com/go/indigo>>. Acesso em: 10 dez. 2008.
- [24] LIN, X. *Active Layout Engine: Algorithms and Applications in Variable Data Printing*. HP Laboratories Palo Alto, CA, USA, 2005. Disponível em: <<http://www.hpl.hp.com/techreports/2005/HPL-2005-6R1.html>>.
- [25] IORIO, A. D. et al. Higher-level layout through topological abstraction. In: *DocEng '08: Proceeding of the eighth ACM symposium on Document engineering*. New York, NY, USA: ACM, 2008. p. 90–99.
- [26] COLLIER, D. *Collier's Rules for Desktop Design and Typography*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1990.
- [27] BALINSKY, H. Y.; WILEY, A. J.; ROBERTS, M. C. Aesthetic measure of alignment and regularity. In: *DocEng '09: Proceedings of the 9th ACM symposium on Document engineering*. New York, NY, USA: ACM, 2009. p. 56–65.

- [28] CHEK, D.; NGO, L.; BYRNE, J. G. Aesthetic measures for screen design. In: *OZCHI '98: Proceedings of the Australasian Conference on Computer Human Interaction*. Washington, DC, USA: IEEE Computer Society, 1998. p. 64.
- [29] BRULS, M.; HUIZING, K.; WIJK, J. J. van. Squarified treemaps. In: *Proc. Joint Eurographics/IEEE TVCG Symp. Visualization, VisSym*. Amsterdam, The Netherlands: IEEE Computer Society, 2000. p. 33–42.
- [30] COFFMAN, E. et al. *Bin packing approximation algorithms: Combinatorial Analysis*. 1998. Disponível em: <<http://citeseer.ist.psu.edu/coffman98bin.html>>. Acesso em: 10 dez. 2007.
- [31] LODI, A.; MARTELLO, S.; VIGO, D. Recent advances on two-dimensional bin packing problems. *Discrete Appl. Math.*, Elsevier Science Publishers B. V., Amsterdam, North Holland, The Netherlands, v. 123, n. 1-3, p. 379–396, 2002.
- [32] DENGLER, E.; FRIEDEL, M.; MARKS, J. Constraint-driven diagram layout. In: *IEEE Symposium on Visual Languages*. Bergen, Norway: IEEE Computer Society, 1993. p. 330–335.
- [33] BATTISTA, G. D. et al. *Graph Drawing: Algorithms for the Visualization of Graphs*. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1998.
- [34] COELLO, C. A. C. A comprehensive survey of evolutionary-based multiobjective optimization techniques. *Knowledge and Information Systems*, v. 1, n. 3, p. 129–156, 1999.
- [35] FLEISCHER, M. Simulated annealing: past, present, and future. In: *WSC '95: Proceedings of the 27th conference on Winter simulation*. Washington, DC, USA: IEEE Computer Society, 1995. p. 155–161.
- [36] JOHARI, R. et al. *Automatic Yellow-Pages Pagination and Layout*. 1997. Disponível em: <<http://citeseer.ist.psu.edu/johari97automatic.html>>. Acesso em: 10 dez. 2008.
- [37] GRÖTSCHEL, M.; LOVÁSZ, L. Combinatorial optimization. MIT Press, Cambridge, MA, USA, p. 1541–1597, 1995.
- [38] MORRISON, M.; BROWNELL, D.; BOUMPHREY, F. *Xml Unleashed*. Indianapolis, IN, USA: Sams, 1999.
- [39] LAMPORT, L. *TEX: A Document Preparation System*. Boston, MA, USA: Addison-Wesley, Reading, Massachusetts, USA, 1986.
- [40] SCHÄFER, C.; PITTMAN, G. *Scribus: Open-Source Desktop Publishing*. United Kingdom: FLES books Ltd., 2009.
- [41] HECKBERT, P. Color image quantization for frame buffer display. *SIGGRAPH Comput. Graph.*, ACM, New York, NY, USA, v. 16, n. 3, p. 297–307, 1982.
- [42] HURST, N.; MARRIOTT, K.; MOULDER, P. Minimum sized text containment shapes. In: *DocEng '06: Proceedings of the 2006 ACM symposium on Document engineering*. New York, NY, USA: ACM, 2006. p. 3–12.
- [43] SKIENA, S. S. *The algorithm design manual*. New York, NY, USA: Springer-Verlag New York, Inc., 1998.

- [44] COLD, S. J. Using really simple syndication (rss) to enhance student research. *SIGITE News/.*, ACM, New York, NY, USA, v. 3, n. 1, p. 6–9, 2006.
- [45] TERRA – RSS. 2008. Disponível em: <<http://www.terra.com.br/rss/>>. Acesso em: 10 dez. 2008.
- [46] FOLHA Online – Informática – Acompanhe notícias da Folha Online no formato RSS. 2008. Disponível em: <<http://www1.folha.uol.com.br/folha/informatica/ult124u16829.shtml>>. Acesso em: 10 dez. 2008.
- [47] HILL, A.; HILL, M. M. *Investigação por Questionário*. 2a.. ed. Lisboa, Portugal: Wook, 2008.
- [48] SETLUR, V. et al. Automatic image retargeting. In: *MUM '05: Proceedings of the 4th international conference on Mobile and ubiquitous multimedia*. New York, NY, USA: ACM, 2005. p. 59–68.

Apêndice A Casos de teste aplicados para a avaliação qualitativa

Nas páginas deste apêndice são apresentados os documentos produzidos pelo algoritmo proposto neste trabalho, e utilizados como caso de teste para a avaliação empírica, descrita no Capítulo 5.



News Document Example



TV estatal da Venezuela veicula gravações de antichavista às vésperas da eleição

Enquanto pesquisas mostram que opositores de Hugo Chávez são favoritos para vencer as eleições regionais marcadas para domingo (23), em alguns dos mais importantes Estados e centros urbanos da Venezuela, a TV estatal entra na campanha, transmitindo grampos de conversas dos adversários do presidente. Segundo o jornal britânico "The Guardian", o principal alvo da estratégia para constrianger os opositores é Manuel Rosales, líder na disputa pela Prefeitura de Maracaibo, a segunda

maior cidade do país. Um anúncio mostra o candidato discutindo a compra de jóias caras, em meio a efeitos sonoros e visuais de anéis e de um relógio de luxo da marca Cartier. Em um outro anúncio, Rosales aparece negociando a compra de gado, tendo ao fundo o som humorístico de mugidos e com imagens de moedas sendo exibidas no vídeo. Até as imagens de um diplomata americano em uma festa de Halloween, vestido de freira e como seios postiços, foram exibidas na TV. Leia mais (18/11/2008 - 14h10)

Novo campo na bacia de Santos vai ampliar produção de gás a partir de dezembro

A oferta de gás natural do país terá acréscimo de 1,5 milhão de metros cúbicos/dia a partir de dezembro, com a entrada em operação do campo de Lagosta, na bacia de Santos, conforme previsão da Petrobras. Segundo o gerente-geral do Ativo Sul da Unidade de Negócios da Bacia de Santos, Luiz Carlos Mendes, boa parte dessa produção será destinada a regiões próximas à Baixada Santista. "Boa parte da produção será destinada à de-

manda daquela área próxima, mas o gás será direcionado a todo o sistema", afirmou Mendes, durante audiência pública relativa ao projeto-piloto do pólo de Tupi, também em Santos. A produção do campo de Lagosta será escoada para o campo de Merluza, que fica ao lado, por meio de um duto de seis quilômetros. A preparação do único poço do campo de Lagosta está na fase final, destacou Mendes. Leia mais (19/11/2008 - 17h09)

STF julgará demarcação da reserva Raposa/Serra do Sol dia 10

O STF (Supremo Tribunal Federal) retomará no próximo dia 10 o julgamento a homologação da terra indígena Raposa/Serra do Sol, em Roraima. O presidente da Suprema Corte, Gilmar Mendes, confirmou a data durante sessão nesta quinta-feira. O julgamento ocorrerá cerca de três meses e meio depois de ser interrompido. No último dia 27 de agosto, o julgamento foi interrompido depois de o relator da ação,

ministro Carlos Ayres Britto, votar pela total retirada dos não-índios da região e pela manutenção da demarcação de forma contínua –sem qualquer interrupção na área delimitada. No entanto, o ministro Carlos Alberto Menezes Direito pediu vista do processo e o julgamento foi adiado. No dia da sessão o ambiente no STF estava tenso: havia representantes de várias etnias indígenas, dos arrozeiros e também políticos. Leia mais (20/11/2008 - 17h41)

Sessão do STF sobre caso Paulo Medina recomeça com discussão de preliminares

Os ministros do STF (Supremo Tribunal Federal) retomaram nesta quinta-feira o julgamento do caso de Paulo Medina, ministro afastado do STJ (Superior Tribunal de Justiça), e de mais quatro pessoas denunciadas de envolvimento com a máfia dos caça-níqueis. A sessão recomeçou com a discussão das chamadas preliminares –questões levantadas pela defesa dos denunciados. Os minis-

tros lamentaram as acusações que colocam em dúvida a credibilidade do Judiciário. "Não posso imaginar um Judiciário que tenha contra si suspeitas que não sejam exauridas exaustivamente investigadas", afirmou o ministro-relator do processo, Cezar Peluso. "Isso para um juiz de carreira é uma coisa extremamente dolorosa. Era preciso que se esclarecesse." Leia mais (20/11/2008 - 11h50)

Acuado, Palmeiras decide reforçar a segurança

O Palmeiras vai reforçar a segurança para evitar que episódios como o da última sexta-feira, quando o técnico Vanderlei Luxemburgo foi agredido por torcedores no aeroporto de Congonhas, voltem a se repetir. O diretor de futebol Genaro Marino afirmou que o assunto será tratado em reunião, hoje. "Em função do que aconteceu, até para deixar o Luxemburgo mais seguro já sugerimos que tivesse [reforço] no desembarque", disse Marino. Leia mais (18/11/2008 - 08h34)

Bandidos tentam roubar carro forte e acabam queimando o dinheiro no Equador

Cerca de US\$ 1 milhão foi queimado no interior de um carro forte que pegou fogo hoje em uma estrada dos Andes equatorianos, depois que ladrões colocaram dinamite perto do tanque de gasolina após não conseguirem abrir o veículo, informou a imprensa local.

Governo tibetano exilado diz que a bola está com a China

O primeiro ministro do governo tibetano no exílio, Samdhong Rinpoche, afirmou nesta terça-feira que as autoridades do Tibete "estão sempre abertas ao diálogo" com Pequim e que agora a "bola está com a China".

Tite tem três candidatos para a vaga de Alex no Inter

Sem o aval da CBF para a liberação do meia Alex, convocado para o amistoso da seleção contra Portugal, o técnico do Internacional, Tite, tenta definir o substituto do atleta para o jogo contra o Chivas, pela Copa Sul-Americana, nesta quarta-feira. A vaga é disputada por Andrezinho, Taison e Daniel Carvalho. Mas, após a derrota do time reserva do Inter para o San-

tos, mesmo que Andrezinho tenha sido poupado como o restante dos titulares, Daniel Carvalho e Taison são os candidatos mais fortes, segundo Tite. "Começo a pensar nisso somente agora. O Andrezinho faz um papel muito parecido com o do D'Alessandro", declarou o treinador do Inter. Leia mais (17/11/2008 - 13h00)

Petrolífera defende ação militar contra pirataria

Recorrer a uma estratégia militar mais agressiva é a única solução para a escalada da pirataria nas costas somalis, segundo Martin Jensen, diretor da companhia Frontline, líder mundial do transporte de petróleo.

Parlamento da Albânia aprova polêmica lei eleitoral

O Parlamento da Albânia aprovou hoje uma polêmica nova lei eleitoral, que provocou uma greve de fome de oito dias de dez deputados que consideram que a reforma discrimina aos partidos minoritários.

Ban expressa a Olmert preocupação com Gaza

O secretário geral da ONU, Ban Ki-moon, ligou nesta terça-feira ao primeiro ministro de Israel, Ehud Olmert, para expressar sua "profunda preocupação" pela situação humanitária em Gaza.

China: 34 trabalhadores presos em mina inundada

Trinta e quatro mineiros estão presos no fundo de uma mina de carvão inundada no centro da China, informou a agência estatal Nova China.

Figura 22: Caso de teste número 1.



News Document Example



Consumidores vão pagar por seqüestro de navios, dizem especialistas

O seqüestro do petroleiro saudita Sirius Star vai pesar no bolso dos consumidores Ocidente, advertem especialistas. Eles antecipam que os cargueiros tomarão agora uma rota mais longa para evitar a ação dos piratas somalis na costa leste da África e ao sul do Iêmen. "Alguém vai ter de pagar o custo dos seqüestros de navios", disse Roger Middleton, especialista do centro de pesquisas Chatham House, de Londres. Para ele, as companhias de transporte

vão rever suas estratégias após o seqüestro do navio da companhia Aramco, carregado com dois milhões de barris de petróleo. O especialista, autor do livro "Pirataria na Somália: Uma Ameaça ao Comércio Internacional", diz que os consumidores da Europa e do norte da África pagarão a conta do seqüestro do Sirius Star, cuja carga está avaliada em US\$ 100 milhões, e de outros navios na região. Leia mais (18/11/2008 - 16h39)

Putin voltará a responder perguntas ao vivo na TV ainda neste ano

O primeiro-ministro da Rússia, Vladimir Putin, cumprirá em dezembro próximo a promessa de voltar a falar com os cidadãos ao vivo pela TV, como fazia quando era presidente, anunciou o gabinete de imprensa do governo nesta terça-feira. "A linha direta acontecerá antes do Ano Novo", destacou Dimitri Peskov, secretário de imprensa de Putin. O secretário de imprensa



O primeiro-ministro Vladimir Putin



Pescadores trabalham durante o entardecer no rio Buranhém, que separa as localidades de Porto Seguro e Arraial D'Ajuda

Cidade histórica, Porto Seguro tem refúgios inspiradores e noite agitada

Considerada Patrimônio Histórico Nacional desde 1973 e reconhecida como Patrimônio Natural da Humanidade pela Unesco em 2000, Porto Seguro é um verdadeiro museu a céu aberto. Marcio Freitas/Folha Imagem Pescadores trabalham durante o entardecer no rio Buranhém, que separa as lo-

calidades de Porto Seguro e Arraial D'Ajuda. Ponto de chegada de Pedro Álvares Cabral, há 508 anos, a cidade não é só axé e badalação. É também um prato cheio para quem quer respirar história e desfrutar dos encantos da natureza. Leia mais (17/11/2008 - 12h07)

Motoristas encontram tráfego bom pelas ruas e avenidas de SP

Os motoristas que trafegam pelas ruas e avenidas da cidade de São Paulo encontram tráfego tranquilo, segundo medições da CET (Companhia de Engenharia de Tráfego). Por volta das 8h50, eram anotados pela companhia 31 km de congestionamento, o equi-

valente a 3,7% dos 835 km monitorados. O trecho mais extenso de lentidão era o do sentido Ayrton Senna da marginal Tietê. Eram anotados pela CET 5,9 km de lentidão pela pista local, a partir do acesso à rodovia dos Bandeirantes. Leia mais (21/11/2008 - 08h48)

TAM estréia hoje rota diária para Orlando, terra da Disney World

A TAM começa a operar hoje a rota São Paulo-Orlando com frequência diária. A cidade, localizada na Flórida (EUA), abriga o maior complexo temático do mundo, a

Disney World, além de parques da Universal e SeaWorld, entre outros. Editora de Arte/Folha Imagem Saiba tudo sobre a Disney Leia mais (21/11/2008 - 11h57)



Diretores do Barclays não receberão bônus em 2008

Os diretores do banco britânico Barclays não receberão bônus neste ano, segundo informou hoje a empresa em comunicado. Ela se une desta maneira a outros bancos, como o americano Goldman Sachs, cujos executivos anunciaram ontem que não receberão esses benefícios. No caso do banco inglês, porém, foi uma espécie de resposta às críticas da decisão de recorrer a in-

vestidores da região do Golfo Pérsico para levar adiante sua ampliação de capital. O Barclays também informou a reeleição de todos os membros da direção —incluindo o executivo-chefe, John Varley, e o chefe da área de banco de investimentos, Bob Diamond— serão submetidos à assembleia geral que será realizada em abril. Leia mais (18/11/2008 - 18h29)

UE e Rússia retomam negociações para novo acordo de associação

A União Europeia (UE) e a Rússia retomam em 2 de dezembro as negociações para a assinatura de um novo acordo de associação que substitua

o vigente desde 1997, afirmou hoje Vladimir Chizhov, embaixador russo em Bruxelas, citado pela agência "Interfax".

Chávez diz que reunião do G20 foi 'decepcionante'

O presidente da Venezuela, Hugo Chávez, qualificou de "decepcionantes" os resultados da reunião do G20 realizada neste fi-

nal de semana em Washington e que tinha como missão encontrar uma saída comum para enfrentar a crise financeira internacional.

Luta contra dengue contará com mais 2.300 militares

Mais de 2.300 militares participarão das tarefas de combate à dengue nos próximos meses em dez estados do país, informou hoje o Ministério da Saúde.

Figura 23: Caso de teste número 2.



News Document Example



FHC admite que errou ao permitir que MPs trancassem pauta do Congresso

O ex-presidente Fernando Henrique Cardoso (PSDB) afirmou nesta terça-feira, em seminário sobre os 20 anos da Constituição Federal na CNI (Confederação Nacional da Indústria), que avalizou um "grave erro" do Congresso Nacional ao permitir que as medidas provisórias tranquem a pauta da Câmara após tramitarem por 45 dias no Legislativo. Ao lado do ministro Tarso Genro (Justiça), FHC disse que o trancamento da pauta permite que o presidente da República bloqueie as atividades legislativas. "Eu concordei com a mudança porque o Congresso achava

que o presidente sufocava o Congresso. Eu disse que estava no fim do mandato e que [a mudança] iria dar ao presidente o poder de bloquear o Congresso. Mas concordei. Isso foi um grave erro do Congresso, eu anui. Mas é preciso tomar cuidado", afirmou. Tarso disse que, como a Constituição de 1988 é "aberta", permite um "certo desequilíbrio" entre os Poderes - o que inclui o excesso de MPs encaminhadas ao Legislativo. Na opinião de Tarso, as medidas provisórias são necessárias para garantir a governabilidade do chefe de Estado. Leia mais (18/11/2008 - 12h06)

Lucro semestral dos seis maiores bancos japoneses cai 58%

O lucro líquido dos seis principais bancos japoneses caiu 58% entre abril e setembro, em comparação com o mesmo período do ano anterior, para 398 bilhões de ienes (US\$ 4,1 bilhões), informou nesta terça-feira a agência de notícias "Kyodo". A crise financeira global afetou intensamente o setor bancário japonês, razão pela qual se prevê que, ao fi-

nal do ano fiscal, que acabará em março, o lucro líquido dos seis bancos se reduzirá pela metade, ficando em 930 bilhões de ienes (US\$ 9,7 bilhões). Os números representariam a primeira queda em quatro anos para uma taxa abaixo de 1 trilhão de ienes (US\$ 10,4 bilhões) do lucro combinado dos seis bancos. Leia mais (18/11/2008 - 10h20)

Justiça de RO condena deputado e oito ex-deputados por cobrarem propina de governador

Acusados de cobrar propina do governador Ivo Cassol (sem partido) em troca de apoio político, um deputado estadual e oito ex-deputados de Rondônia - entre eles um ex-presidente da Assembleia Legislativa - foram condenados nesta segunda-feira pelo Tribunal de Justiça do Estado por formação de quadrilha, concussão (extorsão praticada por funcionário público) e

corrupção passiva. Eles podem recorrer da decisão. As penas variam de seis a 17 anos de prisão. A suposta extorsão foi gravada, no final de 2003, pelo próprio governador em seu escritório e divulgada pelo programa "Fantástico", da Rede Globo, em 2005. Em diferentes conversas, deputados à época apareciam supostamente negociando propina com o governador. Leia mais (17/11/2008 - 21h23)

Galeristas e colecionadores negam crise no Brasil

Se o crash assusta lá fora, o tombo deverá ser menor aqui dentro. Pelo menos é o que esperam galeristas e colecionadores ouvidos pela Folha. Para eles, a arte latino-americana não teve o inchaço registrado no resto do mundo, mantendo seus preços no mesmo patamar ao longo dos anos. "Nós não temos bolha, não temos nenhum Damien Hirst, nenhum Jeff Koons", compara a gale-

rista Luisa Strina. "É uma vantagem porque não vamos numerar de baixar os preços de ninguém." Mas isso não se estende a todos. O galerista Jones Bergamin, diretor da Bolsa de Arte do Rio, acredita que apenas obras mais raras devem manter o preço. "Agora, trabalhos de qualidade média e de arte contemporânea vão ter que dar uma recuada." Leia mais (21/11/2008 - 08h28)

Juiz desiste de inquérito sobre Guerra Civil Espanhola

O juiz espanhol Baltasar Garzón, conhecido por tentar extraditar o ex-ditador chileno Augusto Pinochet, desistiu de abrir um inquérito sobre a morte de aproximadamente 114 mil pessoas durante a Guerra Civil Espanhola, nos anos 30, e durante parte do governo de Francisco Franco. No mês passado, Garzón havia prometido

Gaza: Mahmud Abbas defende manutenção de trégua

O primeiro ministro israelense, Ehud Olmert, anunciou nesta segunda-feira a libertação iminente de 250 prisioneiros palestinos, durante um encontro com o presidente palestino, Mahmud Abbas, que defendeu a manutenção da trégua em Gaza, abalada recentemente por uma onda de violência.

Sarkozy: detenção de "Txeroki" prova colaboração

O presidente francês, Nicolas Sarkozy, assegurou hoje que a detenção do suposto chefe do aparato militar da ETA, Mikel Garikoitz Aspiazu, conhecido como "Txeroki", prova "a excelente colaboração entre França e Espanha na luta contra o terrorismo basco".

Pão de Açúcar é condenado por revistar empregados

O grupo Pão de Açúcar foi condenado pela 3ª Vara do Trabalho de Brasília a pagar indenização de R\$ 1 milhão, revertida para o FAT (Fundo de Amparo ao Trabalhador), por fazer revista em bolsas, sacolas e mochilas de funcionários. O grupo pode recorrer da sentença. A decisão do juiz Francisco Luciano de Azevedo Frota é válida para todas as unidades do grupo no país e prevê ainda que, se a rede voltar a revistar os pertences de seus empregados,

poderá receber multa diária no valor de R\$ 10 mil. "Funcionários prestaram depoimento informando que a empresa fazia a revista e que se sentiam constrangidos. Constatamos que o grupo não fazia revistas íntimas. Mas, ao revistar bolsas, mochilas e sacolas de seus funcionários, a empresa ofende a dignidade do trabalhador, que está protegida pela Constituição", diz a procuradora Ana Cláudia Monteiro. Leia mais (18/11/2008 - 10h21)

Jovem indiano morre engasgado em concurso para comer bolos

Um jovem indiano morreu engasgado durante um concurso no qual os participantes tinham que comer a maior quantidade de bolo possível. Os pais do concorrente denunciaram a falta de pessoal médico no evento. "Este tipo de concurso não deveria ser realizado. Se o concurso acontece, pelo menos deveria ser sob a estrita supervisão de especialistas médicos, para que a his-

tória não se repita", disse o pai da vítima, R. K. Sabharwal, citado pela agência Ians. O jovem Sorab, 22, engenheiro da companhia de telefones celulares Nokia Siemens Networks, morreu ontem quando participava do concurso em uma cantina da localidade de Gurgaon, próxima a Nova Délhi. Leia mais (20/11/2008 - 11h29)

Confirmada identidade de membros do ETA presos

A Guarda Civil espanhola confirmou oficialmente que os dois detidos nesta segunda-feira na França são Garikoitz Aspiazu, conhecido como "Txeroki", e apontado como chefe dos comandos militares do grupo terrorista basco ETA, e a acusada de integrar a organização Leire López Zurutuza.

Equador: US\$ 1 mi é queimado em roubo a carro forte

Cerca de US\$ 1 milhão foi queimado no interior de um carro forte que pegou fogo nesta quinta-feira em uma estrada dos Andes equatorianos, de-

pois que ladrões colocaram dinamite perto do tanque de gasolina após não conseguirem abrir o veículo, informou a imprensa local.

Premiê britânico apóia incentivos fiscais para ativar economia

O primeiro ministro do Reino Unido, Gordon Brown, afirmou hoje que o Reino Unido está disposto a apoiar os incen-

tivos fiscais para ativar a economia, como pediu o Fundo Monetário Internacional (FMI) na cúpula do G20.

Figura 24: Caso de teste número 3.



News Document Example



Roteiro de 24h em Lisboa tem pastel, teleférico, bonde e galeria de arte; leia trecho de guia

Ter pouco tempo para conhecer uma cidade como Lisboa pode ser angustiante. Há muito o que fazer e conhecer na cidade. Diante de tantas opções, vale definir um roteiro e aproveitá-lo ao máximo, mesmo sabendo que qualquer escolha significará deixar de lado outros roteiros igualmente bons. Reprodução Guia seleciona o melhor do design, da arquitetura e do es-

tilo em Lisboa O "Guia Wallpaper Lisboa", editado no Brasil pela Publifolha, sugere seis passeios para curtir Lisboa em 24 horas. As recomendações incluem iniciar a manhã no salão de chá da Pastelaria Mexicana, passear pelo Teleférico às margens do Rio Tejo e visitar a galeria de arte Culturgest. Leia mais (20/11/2008 - 12h12)



Conheça o melhor da arquitetura, design, compras e vida noturna em Lisboa

Preço do petróleo da Opep recua e barril é vendido a US\$ 44,05

O barril de petróleo de referência da Opep (Organização dos Países Exportadores de Petróleo) foi vendido ontem por US\$ 44,05, após cair US\$ 1,83 em relação à cotação anterior, informou nesta sexta-feira o secretariado da organização em Viena (Áustria). Dessa forma, e com novas notícias negativas para o setor automobilístico veiculadas ontem pela imprensa, a cotação do barril

da Opep caiu para o nível mais baixo desde fevereiro de 2005. A queda de preço desde 2 de julho deste ano, quando bateu o recorde histórico de US\$ 140,73, é de 68,7%, e coincide com a sustentada tendência de queda de outros tipos de petróleo, como o Brent e o barril do petróleo leve, de referência para a Europa e para Estados Unidos, respectivamente. Leia mais (21/11/2008 - 10h00)

PSOL faz ato de apoio a Protógenes e De Sanctis em Brasília

O PSOL realiza nesta quarta-feira, a partir das 12h30, uma manifestação em apoio ao delegado federal Protógenes Queiroz, que coordenou a Operação Satiagraha e depois foi afastado do caso pela cúpula da Polícia Federal, e ao juiz Fausto De Sanctis, da 6ª Vara Criminal Federal. O ato ocorrerá em frente ao STF (Supremo Tribunal Federal), no dia em que encerra-se o prazo de defesa do banqueiro Daniel Dantas na Justiça. "Momento em que nos

uniremos perante o STF para pedir a sua condenação e para referendar o trabalho executado pelo delegado Protógenes e pelo juiz De Sanctis", informa o PSOL. De Sanctis foi responsável por decretar as duas prisões temporárias de Dantas durante a Operação Satiagraha. O banqueiro foi solto, nas duas ocasiões, após habeas corpus concedidos pelo presidente do STF (Supremo Tribunal Federal), Gilmar Mendes. Leia mais (18/11/2008 - 11h32)

Perto do retorno, Ronaldo diz que está com 70% de sua forma física

Em fase final de recuperação de uma cirurgia no joelho esquerdo, o atacante Ronaldo afirmou em entrevista que será exibida nesta terça-feira pelo Canal +, da Espanha, que está com 70% de sua forma física e próximo do retorno aos gramados. "Estou com 70%. Quero dizer, muito perto de voltar a jogar. [Minha recuperação] é culpa da paciência do Bruno [Mazziotti, fisio-

terapeuta do jogador]. Ele chegou a me tratar até enquanto eu dormia", revelou o atacante, em trecho antecipado pela edição desta segunda do diário "As". Ronaldo rompeu o tendão patelar do joelho em fevereiro, durante o empate por 1 a 1 do Milan com o Livorno pelo Campeonato Italiano, o que o obrigou a passar por uma cirurgia em Paris, na França. Leia mais (17/11/2008 - 17h30)



Dudu Nobre e Adriana Bombom acusam companhia aérea de racismo e agressão

Adriana Bombom e Dudu nobre acusam companhia aérea de racismo

A modelo e dançarina Adriana Bombom e seu marido, o sambista Dudu Nobre, acusam a companhia aérea American Airlines de agressão e racismo. Bombom afirmou à Folha Online que os problemas começaram ainda nos Estados Unidos, quando ela tentou abrir a porta do banheiro com o avião em terra. Como

o compartimento estava travado, uma comissária se aproximou dela e a ofendeu. "Ela me chamou de estúpida, de ignorante, de tudo", afirmou a dançarina. Renato Rocha Miranda/TV Globo Dudu Nobre e Adriana Bombom acusam companhia aérea de racismo e agressão Leia mais (18/11/2008 - 18h01)

Blog do Josias: Governo prevê PIB e arrecadação menores para 2009

Em meio à recessão das maiores economias do mundo e a crise financeira global, o governo alterou as previsões do PIB (Produto Interno Bruto) e da arrecadação para 2009, informa o blog do Josias. O crescimento do PIB no próximo ano, estimado inicialmente em 4,5%, caiu para 4%, o que implica em menos R\$ 15 bilhões de receitas e gastos públicos para o ano que vem. A perspectiva

do IPCA (Índice de Preços ao Consumidor Amplo) de 4,5% de inflação para 2009 subiu para 5,19%. Da queda na arrecadação, R\$ 10 bilhões referem-se à diminuição na coleta de impostos federais e R\$ 5 bilhões devem-se aos royalties de petróleo. Os novos dados foram remetidos ao Congresso pelo ministro Paulo Bernardo (Planejamento). Leia mais (20/11/2008 - 03h29)

LIVRARIA: Mais de 300 livros em oferta a partir de R\$ 1,99; há viagem, carreira e filhos

Quem gosta de ler deve aproveitar a nova promoção da Livraria da Folha. Mais de 300 livros em oferta por tempo limitado. Entre os títulos disponíveis pelo valor de R\$ 4,90 estão "Cachorro Magro", "Na Hora do Choro" e "Criança

Nota 10", assim como "O Narcotráfico" e "Os Alimentos Transgênicos", da série Folha Explica. Veja todos os livros em promoção ou procure abaixo os livros de seus temas prediletos: Leia mais (19/11/2008 - 23h30)

Itália planeja 1ª conferência sobre crime organizado

O governo italiano, liderado pelo primeiro ministro Silvio Berlusconi, planeja organizar a primeira conferência internacional sobre crime organizado, centrada no tráfico de

drogas, para levar propostas concretas ao Grupo dos Oito (G8, os sete países mais ricos e a Rússia), já que a Itália assume a Presidência do fórum a partir de janeiro.

Maconha pode preservar memória na velhice, diz estudo

Cientistas americanos dizem que substâncias presentes na maconha podem ser benéficas para o cérebro à medida que as pessoas envelhecem, reduzindo índices de inflamação e estimulando a formação de novos neurônios.

Afganistão: combates matam 6 soldados e 5 talibãs

Um violento combate registrado na região de Farah, no oeste do Afeganistão, deixou pelo menos seis soldados afegãos e cinco supostos talibãs mortos, informou hoje à agência EFE uma fonte oficial.

Figura 25: Caso de teste número 4.



News Document Example



Após compra da Nossa Caixa, BB busca aquisições agora no setor privado

O Banco do Brasil não deve recuperar tão cedo a liderança no ranking dos maiores bancos no país nem com uma possível compra de 49% do banco Votorantim, braço financeiro do grupo industrial da família Ermírio de Moraes. Isso porque as atuais regras contábeis impediriam que o BB consolide em seu balanço os ativos do Votorantim, caso fique mesmo como acionista minoritário. Sétima maior instituição do país, o banco Votorantim tem o tamanho certo para devolver a liderança ao BB. Em junho, o banco somava R\$ 73,443 bilhões em

ativos, que, somados aos R\$ 512,4 bilhões de BB e Nossa Caixa, já chegariam a R\$ 585,84 bilhões – mais do que suficientes para passar os R\$ 575 bilhões de Itai e Unibanco. Segundo o professor Ariovaldo dos Santos, da Fipecafi (Fundação Instituto de Pesquisas Contábeis, Atuariais e Financeiras), a única forma de o BB trazer todos os ativos – incluindo crédito e posição de títulos em tesouraria – para o seu balanço é formalizar um acordo em que tenha a gestão compartilhada do novo banco. Leia mais (21/11/2008 - 09h53)

Cientistas observam como buraco negro da Via Láctea devora matéria

Uma equipe internacional de cientistas observou como o buraco negro supermassivo da Via Láctea, com uma massa 4 milhões de vezes maior que a do Sol, segue devorando matéria no centro da galáxia. A descoberta foi anunciada pelo Centro Superior de Pesquisas Científicas da Espanha (Csic). Os pesquisadores detectaram intensos brilhos de Sagittarius A (SgrA), o buraco negro super-

massivo, situado no centro da galáxia, a 26 mil anos-luz da Terra, que revelariam a existência de nuvens de gás desgarradas ao girar em grande velocidade ao redor da formação espacial. Divulgação Conceção artística mostra nuvens de gás em torno de buraco negro; formação continua devorando matéria no centro da galáxia. Leia mais (18/11/2008 - 19h00)



Conceção artística mostra nuvens de gás em torno de buraco negro; pesquisadores afirmam que formação continua devorando matéria no centro da galáxia

Diretor-geral da PF descarta possibilidade de Protógenes reassumir caso contra Dantas

O diretor-geral da Polícia Federal (PF), Luiz Fernando Corrêa, afastou nesta terça-feira a possibilidade do delegado Protógenes Queiroz reassumir a coordenação das investigações relativas à Operação Satiagraha. Corrêa disse ainda que não são "personalizadas" as ações da PF e que Protógenes é "um servidor a mais na Polícia Federal". "A investigação [da Satiagraha] já foi direcionada e tem um novo delegado e uma nova equipe interagindo com o

Ministério Público e com a Justiça produzindo diligências", afirmou Corrêa após cerimônia no Ministério da Justiça, afastando a hipótese de Protógenes retomar a coordenação da operação. Ao ser questionado sobre o futuro de Protógenes, o diretor foi evasivo. "Ele [Protógenes] é um profissional de polícia. Ele é um policial, enquanto ele estiver na PF. Protógenes é um servidor a mais na PF", afirmou Corrêa. Leia mais (18/11/2008 - 11h21)

Defesa Civil permanece em atenção devido a risco de mais chuva no Rio

Depois da chuva que alagou e prejudicou o trânsito no Rio ontem, a cidade amanheceu nesta terça-feira com sol e poucas nuvens. No entanto, ainda há previsão de chuvas rápidas, e a Defesa Civil permanece em estado de atenção. No fim da tarde de segunda-feira (17), uma chuva forte formou pontos de alagamento em vias de toda a cidade, o que gerou congesti-

onamento. O aeroporto Santos Dumont (centro do Rio) fechou por cerca de uma hora, e a Ponte Rio-Niterói, por 20 minutos. Nesta terça, os aeroportos Santos Dumont e internacional Tom Jobim, na Ilha do Governador (zona norte), operam normalmente, segundo a Infraero. O trânsito na cidade também tem poucas retenções, de acordo com a CET-Rio. Leia mais (18/11/2008 - 12h12)

Presidente do Santos admite vender Kléber Pereira após o Nacional

O presidente do Santos, Marcelo Teixeira, admitiu que poderá vender o atacante Kléber Pereira, um dos principais jogadores da equipe nesta temporada, depois do Campeonato Brasileiro. No entanto, ele ressaltou que ainda conversará com o atleta. "Temos que esperar o campeonato acabar e saber qual é o desejo do jogador. Não queremos prejudicar a car-

reira dele. Caso ele queira sair, a proposta precisa ser boa para o clube também", disse o dirigente. Daniel Pereira, irmão e procurador do atacante, afirmou que o centro-avante recebeu uma boa proposta para se transferir para um clube do exterior. Mas, Marcelo Teixeira destacou que o Santos ainda tem contrato com o atleta. Leia mais (20/11/2008 - 11h53)

AGU vai à Justiça para que Abin acompanhe perícia da PF

O advogado geral da União, José Antonio Dias Toffoli, vai pedir ao juiz Ali Mazloum, da 7ª Vara Federal Criminal, que reconsidere o despacho em que proíbe que oficiais da Abin (Agência Brasileira de Inteligência) acompanhem com o Ministério Público a abertura do laço e a verificação de documentos que foram apreendidos

na sede da agência pela Polícia Federal. O juiz proibiu também o acesso da Abin aos exames técnicos em celulares e notebooks recolhidos com arapongas e com o delegado Protógenes Queiroz. Os documentos que estão em posse da PF são considerados secretos e originalmente destinados ao presidente da República. Leia mais (17/11/2008 - 13h24)

Itália planeja 1ª conferência sobre crime organizado

O governo italiano, liderado pelo primeiro ministro Silvio Berlusconi, planeja organizar a primeira conferência internacional sobre crime organizado, centrada no tráfico de drogas, para levar propostas concretas ao Grupo dos Oito (G8, os sete países mais ricos e a Rússia), já que a Itália assume a Presidência do fórum a partir de janeiro.

Colar de 2 mil anos é avaliado em R\$ 1,2 milhões

Uma gargantilha de ouro de cerca de 2 mil anos avaliada em 300 mil libras esterlinas (cerca de R\$ 1,2 milhões) e encontrada por um amador está sendo considerada um dos mais importantes achados do tipo em 50 anos na Grã-Bretanha.



Friedrich Wilhelm Murnau (foto) iria completar seu 120º aniversário neste ano

Paulistanos podem assistir a 12 obras do cineasta alemão Murnau

Divulgação Friedrich Wilhelm Murnau (foto) iria completar 120 anos em 2008. O público paulistano tem até o dia 30 de novembro para assistir aos 12 filmes produzidos pelo cineasta Friedrich Wilhelm Murnau (1888-1931), considerado um dos mais importantes nomes do cinema mudo e do expressionismo alemão. O Centro Cultural Banco do Brasil, na região central da cidade de São Paulo, organizou uma programação que promove uma retrospectiva completa do trabalho do artista, que completaria seu 120º aniversário neste ano. Leia mais (20/11/2008 - 07h42)

Figura 26: Caso de teste número 5.



News Document Example



Polícia investiga hipótese de suicídio em caso de mulher que caiu de prédio em Guarulhos (SP)

A Polícia Civil afirma que trabalha também com a possibilidade de que a operadora de caixa Andréia Cristina Nóbrega, 31, possa ter cometido suicídio após uma discussão com o ex-marido. Ela morreu após cair da janela do terceiro andar do prédio onde morava, em Guarulhos (Grande São Paulo), na terça-feira (18). O filho dela, de 6 anos, caiu na marquete e está internado. Ex-marido de Andréia e pai da criança, Evandro

Correia Filho, 35, é acusado do crime e continua foragido. Um dos indícios de suicídio, diz o delegado Cristiano Engel, está nas cartas que Andréia escreveu à família "em tom de despedida". Os textos estavam no verso de duas fotos achadas no apartamento dela, datadas de fevereiro. "Ela disse para eles não se esqueçam dela e sempre lembrem de seu sorriso", afirmou o delegado. Leia mais (21/11/2008 - 07h56)



Gustavo Patu explica como a carga tributária cresceu tanto no Brasil

PUBLIFOLHA: Livro explica a resistência à reforma tributária e a alta de impostos no Brasil

"Em menos de dez anos a carga tributária deixou o patamar de 25% da renda nacional, no qual havia se mantido por mais de duas décadas, e se aproxima da marca de 36%. A mesma do Reino Unido" - Gustavo Patu, em "A Escalada da Carga Tributária" Divulgação Gustavo Patu explica como a carga tributária cresceu tanto

no Brasil Como é possível que, em menos de uma década, durante tempos de paz, democracia e em plena hegemonia do pensamento liberal, a carga tributária do Brasil tenha crescido tanto? E por que motivos há tanta resistência à realização de uma reforma tributária no país? Leia mais (21/11/2008 - 09h47)

Feriado tem programações especiais nos museus de SP

O município de São Paulo aderiu nesta quinta-feira ao feriado pelo Dia da Consciência Negra. Atracões especiais acontecem por toda a cidade para a comemoração. O Museu do Futebol programou uma visita especial para o feriado com foco na história de aceitação dos negros, mulatos e mestiços no esporte. No Museu Afro Brasil, duas novas exposições têm início neste 20 de novembro, além de ter programada a apresen-

tação de vídeos e performances culturais para comemorar a data. Outra opção é a apresentação de dança prevista para acontecer no Museu da Casa Brasileira, que deve contar com a presença do grupo peruano "Chocolate Riveros y Los Cajones". Também devem acontecer exposições de filmes, como o documentário "Philadelphia - Orixá: Yomonja", que aborda o racismo. Leia mais (20/11/2008 - 11h46)

Inflação em SP fica 0,58% na segunda prévia de novembro, diz Fipe

O IPC (Índice de Preços ao Consumidor) no município de São Paulo registrou uma alta de 0,58% na segunda quadrissemana de novembro -30 dias até 15/11-, ligeira correção para cima em relação ao índice visto na abertura do mês, 0,57%. Os dados foram divulgados nesta terça-feira pela Fipe (Fundação Instituto de Pesquisas Econômicas). Entenda a diferença entre os principais índices de inflação A categoria Alimentação registrou a maior variação percentual de preços, 70%, mas a contribuição para a composição final do indicador foi de 27,28%. Já a categoria Habitação teve uma variação ligeiramente menor nos preços, 0,69%, mas teve maior peso na composição do índice. Leia mais (18/11/2008 - 08h46)



Com os filhos Zahara (à dir.) e Shiloh, Brad Pitt chega ao hospital onde está Angelina

Shiloh quer ser chamada de John, diz Brad Pitt

Em entrevista a Oprah Winfrey, Brad Pitt revelou uma mania da filha de dois anos e meio, Shiloh. Lionel Cironneau/AP Brad Pitt com Zahara (à dir.) e Shiloh; "Angie está sempre no comando", contou o ator "Ela só quer ser chamada de John. John ou Peter. É algo do Peter Pan", disse. Leia mais (18/11/2008 - 21h11)

Sarcófagos achados no palácio de Herodes são exibidos

Três sarcófagos com cerca de 2 mil anos descobertos em escavações no palácio de Herodes, rei da Judéia (37 a.C. e 4 a.C.), foram exibidos na Universidade Hebraica, em

Jerusalém, informa a agência Reuters. Os pesquisadores israelenses acreditam que em duas das tumbas teriam sido enterradas a mulher e a nora de Herodes.

Cientistas observam buraco negro da Via Láctea

Uma equipe internacional de cientistas observou como o buraco negro supermassivo da Via Láctea, com uma massa 4 milhões de vezes maior que a do Sol,

segue devorando material no centro da galáxia, informou nesta terça o Centro Superior de Pesquisas Científicas da Espanha (Csic).

Gene pode explicar dependência de cocaína

Parte da população mundial pode ser mais vulnerável ao vício de cocaína devido a uma variação genética, de acordo com uma nova pesquisa feita por cientistas de diversos países.

Figura 27: Caso de teste número 6.



News Document Example



Maranhão diz que espera publicação de decisão do TSE para assumir governo da Paraíba

O senador José Maranhão (PMDB-PB) disse nesta sexta-feira que aguarda a publicação do acórdão do TSE (Tribunal Superior Eleitoral) para renunciar ao cargo e assumir o governo da Paraíba. Sorridente, Maranhão afirmou que a decisão chegou "com grande retardo", mas confirma a vitória que ele disse merecer. Ele também negou que tenha sonogado informações sobre sua evolução patrimonial. "Todo meu patrimônio está decla-

rado e claramente demonstrado. Todos os homens públicos correm esses riscos [de supostas falsas denúncias]", afirmou Maranhão, que antes de ir ao seu gabinete se reuniu com os advogados de defesa. O senador se referiu às informações, que vieram à tona no ano passado, de que ele não informou à Receita Federal sobre sua evolução patrimonial, deixando de declarar 28.290 cabeças de gado. Leia mais (21/11/2008 - 13h42)

Bolsa de NY cai mais de 5% e fecha no menor nível em cinco anos

A Bolsa de Nova York fechou em queda nesta quinta-feira, que a levou ao seu nível mais baixo em cinco anos e meio, afetada pela degradação da economia norte-americana e pela ausência de uma resposta política à crise na indústria automobilística. O índice Dow Jones Industrial Average, da Nyse (Bolsa de Valores de Nova York, na sigla em inglês), recuou 5,56%, aos 7.552,29 pontos - no seu nível mais baixo

desde março de 2003. Em um pregão volátil, o Dow Jones chegou a descer 6,1% e a subir 2,37%. O índice Nasdaq, de elevado componente tecnológico, perdeu 5,07%, aos 1.316,12 pontos. Já o S&P 500 caiu 6,71%, a 752,44 pontos - nível não registrado desde abril de 1997. "É muito perigoso. O mercado está cauteloso", disse Lindsey Piegza, da FTN Financial. Leia mais (20/11/2008 - 20h26)

STJ veta alta em plano de saúde para idosos associados da Unimed Natal; ouça

Na semana passada, o STJ (Superior Tribunal de Justiça) proibiu o aumento das mensalidades de planos de saúde por faixa etária para quem tem 60 anos ou mais. O entendimento vale apenas para os associados da Unimed Natal. Maria Inês Dolci, coordenadora institucional da Pro Teste, colunista da Folha e responsável por um blog da

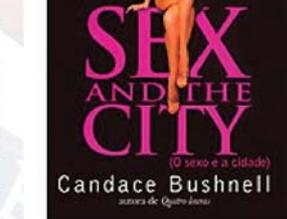
Folha Online, diz que a empresa informou aos seus clientes de planos antigos que a partir de 2004 iria aumentar em 100% e 200% as mensalidades para quem completasse 60 e 70 anos, respectivamente. Ouça outros podcasts com a participação da advogada. Leia mais (17/11/2008 - 12h00)

EUA: após incêndios, ator encontra casa destruída

Obrigado a o ator Christopher Lloyd devido aos incêndios que encontrou a residência de US\$ 11 milhões completamente destruída, informou a rede CNN.

LIVRARIA: Veja como foi o encontro de Carrie e Mr. Big em "Sex and The City"

A narrativa do primeiro encontro de Carrie e Mr. Big está no livro "Sex and The City", que traz as histórias originais que deram origem à série de TV e ao longa-metragem de mesmo nome. Reprodução Livro traz as histórias originais



Livro traz as histórias picantes de Carrie, Samantha, Miranda e Charlotte

Posição de carrinho 'pode afetar desenvolvimento do bebê'

Bebês e crianças que são levadas em carrinhos que as deixam de costas para os pais podem ter seu desenvolvimento prejudicado, sugere um estudo da Universidade de Dundee, na Escócia.

Conselho Eleitoral marca eleições no Equador

As eleições gerais no Equador ocorrerão em 26 de abril de 2009, informou nesta terça-feira o presidente do Conselho Nacional Eleitoral (CNE), Omar Simons.

Figura 28: Caso de teste número 7.



News Document Example



Dilma diz que investir no pré-sal é questão anticíclica para Brasil

A ministra Dilma Rousseff (Casa Civil) disse nesta sexta-feira que os investimentos na camada do pré-sal têm caráter anticíclico contra a crise mundial e que manter o volume elevado de recursos na exploração dessa área permite que o Brasil aproveite as "oportunidades de crescimento" no futuro. "A descoberta de hoje, para nós, é muito importante porque configura um dos fatores anticrise no Brasil. A decisão do governo de manter os investimentos no PAC [Programa de Aceleração do Crescimento] e no pré-sal é importante porque são questões eminen-

temente anticíclicas", disse ela, numa referência ao anúncio de novas reservas na camada do pré-sal no Espírito Santo, divulgado hoje pela Petrobras. Ainda comentando as descobertas de hoje, a ministra afirmou: "estamos caminhando na direção de constituir o pré-sal como uma das opções para os próximos anos, para as próximas décadas. Por isso, o governo mantém a decisão de investir no pré-sal e acha que o investimento dessa qualidade é algo fundamental como oportunidade de crescimento para o país". Leia mais (21/11/2008 - 13h54)

Chip's Burger lança sanduíche para acompanhar balada em SP

Mauro Holanda/Divulgação
Sanduíche de pernil do Chip's Burger é servido às sextas e sábados, na madrugada. A casa, que tem a intenção de recriar uma atmosfera norte-americana, serve em uma das vias mais agitadas de Santana, próxima à avenida Braz Leme, onde há grande concentração de bares e restaurantes, está a - 20h44)



Sanduíche de pernil do Chip's Burger é servido às sextas e sábados, na madrugada

Milícia governista e rebeldes enfrentam-se no Congo um dia após início de retirada

Os rebeldes começaram esta manhã nas cidades de Laurent Nkunda e Nyongera, travaram combates 5 km ao norte de nesta quinta-feira Kiwanja", afirmou com milícias favoráveis ao governo na região leste da República Democrática do Congo (RDC, antigo Zaire), um dia depois de um recuo rebelde na linha de combate. "Os enfrentamentos

começaram esta manhã nas cidades de Katoro e Nyongera, 5 km ao norte de Kiwanja", afirmou uma fonte dos serviços de segurança, que pediu anonimato. Kiwanja fica 75 km ao norte de Goma, a capital da província de Kivu Norte. Leia mais (20/11/2008 - 11h19)

Emenda pode cortar verba de partido neonazista alemão

Os ministros de Interior extrema direita, por meio do governo alemão e os Estados federados pretendem cortar o financiamento Estal aos partidos nazistas de uma emenda na Constituição que vetaria o repasse dos fundos.

Bomba mata líder de família mafiosa em Tel Aviv

Yacov Alperon, considerado o chefe da maior família ligada ao crime organizado em Israel, foi assassinado nesta segunda-feira, quando uma bomba explodiu no carro em que estava, em uma das ruas principais de Tel Aviv.

Ban expressa a Olmert preocupado com Gaza

O secretário geral da ONU, Ban Ki-moon, ligou nesta terça-feira ao primeiro ministro de Israel, Ehud Olmert, para expressar sua "profunda preocupação" pela situação humanitária em Gaza.

Figura 29: Caso de teste número 8.



News Document Example



Jobim se reúne com Lula para discutir Plano Nacional de Defesa

O governo federal prepara uma série de discussões ainda neste ano sobre o Plano Nacional de Defesa. A medida, segundo o ministro Nelson Jobim (Defesa), vai reverter a obrigatoriedade do serviço militar no país, o fortalecimento da indústria bélica e a recolocação dos militares na sociedade, entre outros pontos. A agenda para o plano começa na terça-feira, quando ele apresentará a proposta a parlamentares da Câmara e do Senado.

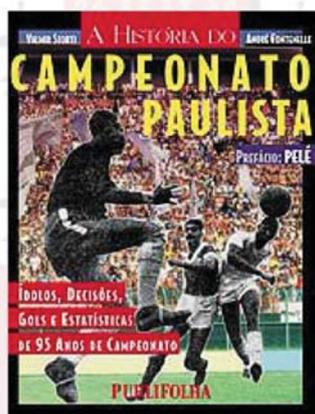
Na quinta-feira, Jobim tem uma reunião no Palácio do Planalto com o presidente Luiz Inácio Lula da Silva e ministros Mangabeira Unger (Assuntos Estratégicos), Dilma Rousseff (Casa Civil) e Franklin Martins (Comunicação) para concluir o plano. Em seguida, Jobim convocará o Conselho de Defesa Nacional. O ministro estima aprovar o plano até 15 de dezembro. Leia mais (20/11/2008 - 18h17)

Índia afirma que G20 superou G7 na resolução de problemas econômicos

O ministro de Finanças indiano, P. Chidambaram, disse nesta terça-feira que o G20 (grupo que reúne os países mais ricos e os principais emergentes) transformou-se no fórum mundial mais importante para resolver problemas econômicos. A declaração indiana segue a linha adotada pelo presidente Luiz Inácio Lula da Silva, mais (18/11/2008 - 13h50)

Polícia prende dez pessoas em tentativa de assalto a transportadora em Canoas (RS)

Policiais da Brigada Militar prenderam dez pessoas na madrugada desta sexta-feira após uma tentativa de assalto à transportadora Transpaulo, no bairro de Matias Velho, em Canoas (RS). De acordo com a polícia, um grupo fortemente armado invadiu a empresa por volta da 0h30 e ren-



Livro traz ídolos, decisões, Gols e estatísticas de 95 anos de Campeonato

LIVRARIA: Mais de 100 livros de R\$ 1,99 a R\$ 9,90; há viagens, filhos, carreira e receitas

Veja abaixo completa a lista de mais de 350 títulos no site. Veja ainda uma lista de promoções por R\$ 1,99 a R\$ 9,90. A lista é atualizada diariamente com novas promoções. Confira também a lista

AIEA confirma que há urânio em instalação síria

A Agência Internacional de Energia Atômica (AIEA) confirmou hoje em Viena que foram encontrados rastros de urânio em uma suposta instalação militar na Síria, bombardeada ano passado pela aviação israelense e onde os Estados Unidos asseguram que ia ser construído um reator nuclear.

Figura 30: Caso de teste número 9.



News Document Example



Novela: Luca convida Dani para trabalhar no Nonna Mia em “Água na Boca”

A programação está sujeita a alteração devido à edição dos capítulos e é de responsabilidade da emissora Água na Boca (Band, 20h10) Keiko (Eda Nagayama) decide enfrentar Artur (Peterlongo/João Signorelli) e conta com a ajuda de Kim (Alexandre Moreno). Françoise (Jacqueline Laurence) e Jean-Paul (Mário César Camargo) chegam ao Mercado. Zeca (Joaquim Lopes) encara Françoise. Miquelina (Ana Cecília) diz para Severino (Claudio Jaborandy) que irá filmar o 'chamego' com Bertinho (Betito Tavares). Luca (Caetano O'Maihlán) pede para as Mamas cuidarem de Allegra (Karen Tavares). Jean-Paul conta para Zeca que os Cassou-

let perderam tudo. Érika (Marisol Ribeiro) conta para Alex que Luca tem uma filha. Ele quer saber se ela está pronta para trabalhar no Paris. Dani sai em busca de emprego, mas ninguém a aceita. Alex propõe a Remi (Marco Antônio Pâmio) que volte a trabalhar no Paris. Dani descobre que Alex divulgou o escândalo de sua família no jornal. Alex convoca uma reunião com os funcionários do Paris. Luca vai atrás de Gina (Miranda Kassin) para saber porque ela escondeu Allegra por tantos anos. Bertinho chega ao Baião de Dois. Alex quer recontratar os funcionários do Paris e convida Érika para ser chef. Keiko mostra a foto de Artur para Kim. Severino vai atrás de Miquelina e Berti-

nho, que estão no quarto. Os dois fingem que estão transando. Érika finge surpresa a proposta de Alex. Dani entra no Bistrô e parte para cima dele. Severino (Claudio Jaborandy) vê Miquelina (Ana Cecília) e Bertinho (Betito Tavares) na cama. Ele sai magoado e Miquelina se arrende. Dani (Rosanne Mulholland) diz para Alex (Alexandre Barros) que não vai desistir de sua vida. Érika (Marisol Ribeiro) leva Dani para a sua casa. Luca (Caetano O'Maihlán) convida Dani para trabalhar no Nonna Mia. Ela promete pensar. Allegra (Karen Tavares) tem um pesadelo e vai dormir com Luca e Érika, que fica uma fera. Leia mais (18/11/2008 - 00h16)



"Em reunião com Barack Obama, McCain prometeu 'unir forças' para reformas" <http://www1.folha.uol.com.br/folha/mundo/ult9pgs6628.shtml>

Após derrota nas eleições, McCain volta ao Senado

O ex-candidato republicano à Casa Branca, John McCain, senador do Arizona, voltou nesta terça-feira, 12.nov.08/Efe Em reunião com Barack Obama, McCain prometeu "unir forças" para reformas O líder da minoria republicana, o senador Mitch McConnell, afirmou estar "contente" com o retorno do ex-candidato à Presidência. Leia mais (18/11/2008 - 18h00)

Ruas e avenidas do Rio apresentam tráfego intenso nesta terça-feira

As principais ruas e avenidas no Rio apresentam tráfego intenso na tarde desta terça-feira, segundo a prefeitura. Por volta das 17h30, havia lentidão no centro e nas regiões norte e sul da cidade. O motorista reduzia a velocidade no centro, na avenida Perimetral, e na altura da Praça Mauá, como na altura da

rua Professor Pereira Reis, ambos os trechos no sentido avenida Brasil; e na rua Frei Caneca, altura da rua do Santana. Na zona norte, o acesso às vias estava complicado no entorno da praça da Bandeira, no sentido Tijuca; e na avenida Brasil, altura da rua Prefeito Olímpio de Melo, na pista lateral, no sentido zona oeste Leia mais (18/11/2008 - 17h40)

PUBLIFOLHA: Veja obras da série "Folha Explica" em promoção; títulos a partir de R\$ 3,90

Quem gosta de ler deve aproveitar a promoção da Publifolha. Veja a lista de livros da série "Folha Explica" em oferta por tempo limitado a partir de R\$ 3,90 e abasteça a sua biblioteca. O preço normal dos livros varia entre R\$ 17,90 e R\$ 20,90. Procure livros de outros temas AQUI. * Leia mais (18/11/2008 - 13h11)

ONU: Rice participa de combate a violência contra mulher

A secretária de Estado americana, Condoleezza Rice, assinou uma campanha promovida pela ONU para pressionar os governos domundo a acabar com a violência contra as mulheres.

Figura 31: Caso de teste número 10.



News Document Example



TV estatal da Venezuela veicula gravações de antichavista às vésperas da eleição

Enquanto pesquisas mostram que opositores de Hugo Chávez são favoritos para vencer as eleições regionais marcadas para domingo (23), em alguns dos mais importantes Estados e centros urbanos da Venezuela, a TV estatal entra na campanha, transmitindo grampos de conversas dos adversários do presidente. Segundo o jornal britânico "The Guardian", o principal alvo da estratégia para constrianger os opositores é Manuel Rosales, líder na disputa pela Prefeitura de Maracaibo, a segunda

maior cidade do país. Um anúncio mostra o candidato discutindo a compra de jóias caras, em meio a efeitos sonoros e visuais de anéis e de um relógio de luxo da marca Cartier. Em um outro anúncio, Rosales aparece negociando a compra de gado, tendo ao fundo o som humorístico de mugidos e com imagens de moedas sendo exibidas no vídeo. Até as imagens de um diplomata americano em uma festa de Halloween, vestido de freira e como seios postiços, foram exibidas na TV. Leia mais (18/11/2008 - 14h10)

Novo campo na bacia de Santos vai ampliar produção de gás a partir de dezembro

A oferta de gás natural do país terá acréscimo de 1,5 milhão de metros cúbicos/dia a partir de dezembro, com a entrada em operação do campo de Lagosta, na bacia de Santos, conforme previsão da Petrobras. Segundo o gerente-geral do Ativo Sul da Unidade de Negócios da Bacia de Santos, Luiz Carlos Mendes, boa parte dessa produção será destinada a regiões próximas à Baixada Santista. "Boa parte da produção será destinada à de-

manda daquela área próxima, mas o gás será direcionado a todo o sistema", afirmou Mendes, durante audiência pública relativa ao projeto-piloto do pólo de Tupi, também em Santos. A produção do campo de Lagosta será escoada para o campo de Merluza, que fica ao lado, por meio de um duto de seis quilômetros. A preparação do único poço do campo de Lagosta está na fase final, destacou Mendes. Leia mais (19/11/2008 - 17h09)

STF julgará demarcação da reserva Raposa/Serra do Sol dia 10

O STF (Supremo Tribunal Federal) retomará no próximo dia 10 o julgamento a homologação da terra indígena Raposa/Serra do Sol, em Roraima. O presidente da Suprema Corte, Gilmar Mendes, confirmou a data durante sessão nesta quinta-feira. O julgamento ocorrerá cerca de três meses e meio depois de ser interrompido. No último dia 27 de agosto, o julgamento foi interrompido depois de o relator da ação,

ministro Carlos Ayres Britto, votar pela total retirada dos não-índios da região e pela manutenção da demarcação de forma contínua –sem qualquer interrupção na área delimitada. No entanto, o ministro Carlos Alberto Menezes Direito pediu vista do processo e o julgamento foi adiado. No dia da sessão o ambiente no STF estava tenso: havia representantes de várias etnias indígenas, dos arrozeiros e também políticos. Leia mais (20/11/2008 - 17h41)

Sessão do STF sobre caso Paulo Medina recomeça com discussão de preliminares

Os ministros do STF (Supremo Tribunal Federal) retomaram nesta quinta-feira o julgamento do caso de Paulo Medina, ministro afastado do STJ (Superior Tribunal de Justiça), e de mais quatro pessoas denunciadas de envolvimento com a máfia dos caça-níqueis. A sessão recomeçou com a discussão das chamadas preliminares –questões levantadas pela defesa dos denunciados. Os minis-

tros lamentaram as acusações que colocam em dúvida a credibilidade do Judiciário. "Não posso imaginar um Judiciário que tenha contra si suspeitas que não sejam exauridas exaustivamente investigadas", afirmou o ministro-relator do processo, Cezar Peluso. "Isso para um juiz de carreira é uma coisa extremamente dolorosa. Era preciso que se esclarecesse." Leia mais (20/11/2008 - 11h50)

Acuado, Palmeiras decide reforçar a segurança

O Palmeiras vai reforçar a segurança para evitar que episódios como o da última sexta-feira, quando o técnico Vanderlei Luxemburgo foi agredido por torcedores no aeroporto de Congonhas, voltem a se repetir. O diretor de futebol Genaro Marino afirmou que o assunto será tratado em reunião, hoje. "Em função do que aconteceu, até para deixar o Luxemburgo mais seguro já sugerimos que tivesse [reforço] no desembarque", disse Marino. Leia mais (18/11/2008 - 08h34)

Bandidos tentam roubar carro forte e acabam queimando o dinheiro no Equador

Cerca de US\$ 1 milhão foi queimado no interior de um carro forte que pegou fogo hoje em uma estrada dos Andes equatorianos, depois que ladrões colocaram dinamite perto do tanque de gasolina após não conseguirem abrir o veículo, informou a imprensa local.

Governo tibetano exilado diz que a bola está com a China

O primeiro ministro do governo tibetano no exílio, Samdhong Rinpoche, afirmou nesta terça-feira que as autoridades do Tibete "estão sempre abertas ao diálogo" com Pequim e que agora a "bola está com a China".

Tite tem três candidatos para a vaga de Alex no Inter

Sem o aval da CBF para a liberação do meia Alex, convocado para o amistoso da seleção contra Portugal, o técnico do Internacional, Tite, tenta definir o substituto do atleta para o jogo contra o Chivas, pela Copa Sul-Americana, nesta quarta-feira. A vaga é disputada por Andrezinho, Taison e Daniel Carvalho. Mas, após a derrota do time reserva do Inter para o San-

tos, mesmo que Andrezinho tenha sido poupado como o restante dos titulares, Daniel Carvalho e Taison são os candidatos mais fortes, segundo Tite. "Começo a pensar nisso somente agora. O Andrezinho faz um papel muito parecido com o do D'Alessandro", declarou o treinador do Inter. Leia mais (17/11/2008 - 13h00)

Petrolífera defende ação militar contra pirataria

Recorrer a uma estratégia militar mais agressiva é a única solução para a escalada da pirataria nas costas somalis, segundo Martin Jensen, diretor da companhia Frontline, líder mundial do transporte de petróleo.

Parlamento da Albânia aprova polêmica lei eleitoral

O Parlamento da Albânia aprovou hoje uma polêmica nova lei eleitoral, que provocou uma greve de fome de oito dias de dez deputados que consideram que a reforma discrimina aos partidos minoritários.

Ban expressa a Olmert preocupação com Gaza

O secretário geral da ONU, Ban Ki-moon, ligou nesta terça-feira ao primeiro ministro de Israel, Ehud Ol-

mert, para expressar sua "profunda preocupação" pela situação humanitária em Gaza.

China: 34 trabalhadores presos em mina inundada

Trinta e quatro mineiros estão presos no fundo de uma mina de carvão inundada no centro da China, informou a agência estatal Nova China.

Figura 32: Caso de teste número 11.



News Document Example



Consumidores vão pagar por seqüestro de navios, dizem especialistas

O seqüestro do petroleiro saudita Sirius Star vai pesar no bolso dos consumidores do Ocidente, advertem especialistas. Eles antecipam que os cargueiros tomarão agora uma rota mais longa para evitar a ação dos piratas somalis na costa leste da África e ao sul do Iêmen. "Alguém vai ter de pagar o custo dos seqüestros de navios", disse Roger Middleton, especialista do centro de pesquisas Chatham House, de Londres. Para ele, as companhias de transporte vão

rever suas estratégias após o seqüestro do navio da companhia Aramco carregado com dois milhões de barris de petróleo. O especialista, autor do livro "Pirataria na Somália: Uma Ameaça ao Comércio Internacional", diz que os consumidores da Europa e do norte da África pagarão a conta do seqüestro do Sirius Star, cuja carga está avaliada em US\$ 100 milhões, e de outros navios na região. Leia mais (18/11/2008 - 16h39)

Diretores do Barclays não receberão bônus em 2008

Os diretores do banco britânico Barclays não receberão bônus neste ano, segundo informou hoje a empresa em comunicado. Ela se une desta maneira a outros bancos, como o americano Goldman Sachs, cujos executivos anunciaram ontem que não receberão esses benefícios. No caso do banco inglês, porém, foi uma espécie de resposta às críticas da decisão de re-

correr a investidores da região do Golfo Pérsico para levar adiante sua ampliação de capital. O Barclays também informou a reeleição de todos os membros da direção—incluindo o executivo-chefe, John Varley, e o chefe da área de banco de investimentos, Bob Diamond—seção submetidos à assembleia geral que será realizada em abril. Leia mais (18/11/2008 - 18h29)

Luta contra dengue contará com mais 2.300 militares

Mais de 2.300 militares participarão das tarefas de combate à dengue nos próximos meses em dez estados do país, informou hoje o Ministério da Saúde.

UE e Rússia retomarão negociações para novo acordo de associação

A União Europeia (UE) e a Rússia retomarão em 2 de dezembro as negociações para a assinatura de um novo acordo de associação que substitua o vigente desde 1997, afirmou hoje Vladimir Chizhov, embaixador russo em Brasília, citado pela agência "Interfax".

Motoristas encontram tráfego bom pelas ruas e avenidas de SP

Os motoristas que trafegam pelas ruas e avenidas da cidade de São Paulo encontram tráfego tranquilo, segundo medições da CET (Companhia de Engenharia de Tráfego). Por volta das 8h50, eram anotados pela companhia 31 km de congestionamento, o equivalente a 3,7% dos 835 km monitorados. O trecho mais extenso de lentidão era o do sentido Ayrton Senna da marginal Tietê. Eram anotados pela CET 5,9 km de lentidão pela pista local, a partir do acesso à rodovia dos Bandeirantes. Leia mais (21/11/2008 - 08h48)

TAM estréia hoje rota diária para Orlando, terra da Disney World

A TAM começa a operar hoje a rota São Paulo-Orlando com frequência diária. A cidade, localizada na Flórida (EUA), abriga o maior complexo temático do mundo, a Disney World,

além de parques da Universal e SeaWorld, entre outros. Editora de Arte/Folha Imagem Saiba tudo sobre a Disney Leia mais (21/11/2008 - 11h57)



Putin voltará a responder perguntas ao vivo na TV ainda neste ano

O primeiro-ministro da Rússia, Vladimir Putin, cumprirá em dezembro próximo a promessa de voltar a falar com os cidadãos ao vivo pela TV, como fazia quando era presidente, anunciou o gabinete de imprensa do governo nesta terça-feira. "A linha direta acontecerá antes do Ano Novo", destacou Dimitri Peskov, secretário de imprensa de Putin. O secretário de imprensa ex-

plicou que Putin responderá às perguntas dos cidadãos "tanto na qualidade de primeiro-ministro quanto na de líder do partido Rússia Unida" e acrescentou que o "leque de perguntas será bastante amplo". Destacou que a "linha direta" será transmitida pela Corporação Estatal de Rádio e Televisão. AP O primeiro-ministro Vladimir Putin Leia mais (18/11/2008 - 13h22)

Chávez diz que reunião do G20 foi 'decepcionante'

O presidente da Venezuela, Hugo Chávez, qualificou de "decepcionantes" os resultados da reunião do G20 realizada neste final de semana

em Washington e que tinha como missão encontrar uma saída comum para enfrentar a crise financeira internacional.



Pescadores trabalham durante o entardecer no rio Buranhém, que separa as localidades de Porto Seguro e Arraial D'Ajuda



O primeiro-ministro Vladimir Putin

Cidade histórica, Porto Seguro tem refúgios inspiradores e noite agitada

Considerada Patrimônio Histórico Nacional desde 1973 e reconhecida como Patrimônio Natural da Humanidade pela Unesco em 2000, Porto Seguro é um verdadeiro museu a céu aberto. Marcio Freitas/Folha Imagem Pescadores trabalham durante o entardecer no rio Buranhém, que separa as

localidades de Porto Seguro e Arraial D'Ajuda. Ponto de chegada de Pedro Álvares Cabral, há 508 anos, a cidade não é só axé e badalação. É também um prato cheio para quem quer respirar história e desfrutar dos encantos da natureza. Leia mais (17/11/2008 - 12h07)

Figura 33: Caso de teste número 12.



News Document Example



FHC admite que errou ao permitir que MPs trancassem pauta do Congresso

O ex-presidente Fernando Henrique Cardoso (PSDB) afirmou nesta terça-feira, em seminário sobre os 20 anos da Constituição Federal na CNI (Confederação Nacional da Indústria), que avalizou um "grave erro" do Congresso Nacional ao permitir que as medidas provisórias tranquem a pauta da Câmara após tramitarem por 45 dias no Legislativo. Ao lado do ministro Tarso Genro (Justiça), FHC disse que o trancamento da pauta permite que o presidente da República bloqueie as atividades legislativas. "Eu concordei com a mudança porque o Congresso achava que o presi-

dente sufocava o Congresso. Eu disse que estava no fim do mandato e que [a mudança] iria dar ao presidente o poder de bloquear o Congresso. Mas concordei. Isso foi um grave erro do Congresso, eu anui. Mas é preciso tomar cuidado", afirmou. Tarso disse que, como a Constituição de 1988 é "aberta", permite um "certo desequilíbrio" entre os Poderes - o que inclui o excesso de MPs encaminhadas ao Legislativo. Na opinião de Tarso, as medidas provisórias são necessárias para garantir a governabilidade do chefe de Estado. Leia mais (18/11/2008 - 12h06)

Juiz desiste de inquérito sobre Guerra Civil Espanhola

O juiz espanhol Baltasar Garzón, conhecido por tentar extraditar o ex-ditador chileno Augusto Pinochet, desistiu de abrir um inquérito sobre a morte de aproximadamente 114 mil pessoas durante a Guerra Civil Espanhola, nos anos 30, e durante parte do governo de Francisco Franco. No mês passado, Garzón havia prometido

Gaza: Mahmud Abbas defende manutenção de trégua

O primeiro ministro israelense, Ehud Olmert, anunciou nesta segunda-feira a libertação iminente de 250 prisioneiros palestinos, durante um encontro com o presidente palestino, Mahmud Abbas, que defendeu a manutenção da trégua em Gaza, abalada recentemente por uma onda de violência.

Lucro semestral dos seis maiores bancos japoneses cai 58%

O lucro líquido dos seis principais bancos japoneses caiu 58% entre abril e setembro, em comparação com o mesmo período do ano anterior, para 398 bilhões de ienes (US\$ 4,1 bilhões), informou nesta terça-feira a agência de notícias "Kyodo". A crise financeira global afetou intensamente o setor bancário japonês, razão

pela qual se prevê que, ao final do ano fiscal, que acabará em março, o lucro líquido dos seis bancos se reduzirá pela metade, ficando em 930 bilhões de ienes (US\$ 9,7 bilhões). Os números representariam a primeira queda em quatro anos para uma taxa abaixo de 1 trilhão de ienes (US\$ 10,4 bilhões) do lucro combinado dos seis bancos. Leia mais (18/11/2008 - 10h20)

Pão de Açúcar é condenado por revistar empregados

O grupo Pão de Açúcar foi condenado pela 3ª Vara do Trabalho de Brasília a pagar indenização de R\$ 1 milhão, revertida para o FAT (Fundo de Amparo ao Trabalhador), por fazer revista em bolsas, sacolas e mochilas de funcionários. O grupo pode recorrer da sentença. A decisão do juiz Francisco Luciano de Azevedo Frota é válida para todas as unidades do grupo no país e prevê ainda que, se a rede voltar a revistar os pertences de seus empregados,

poderá receber multa diária no valor de R\$ 10 mil. "Funcionários prestaram depoimento informando que a empresa fazia a revista e que se sentiam constrangidos. Constatamos que o grupo não fazia revistas íntimas. Mas, ao revistar bolsas, mochilas e sacolas de seus funcionários, a empresa ofende a dignidade do trabalhador, que está protegida pela Constituição", diz a procuradora Ana Claudia Monteiro. Leia mais (18/11/2008 - 10h21)

Justiça de RO condena deputado e oito ex-deputados por cobrarem propina de governador

Acusados de cobrar propina do governador Ivo Cassol (sem partido) em troca de apoio político, um deputado estadual e oito ex-deputados de Rondônia - entre eles um ex-presidente da Assembleia Legislativa - foram condenados nesta segunda-feira pelo Tribunal de Justiça do Estado por formação de quadrilha, concussão (extorsão praticada por funcionário público) e corrupção passiva. Eles

podem recorrer da decisão. As penas variam de seis a 17 anos de prisão. A suposta extorsão foi gravada, no final de 2003, pelo próprio governador em seu escritório e divulgada pelo programa "Fantástico", da Rede Globo, em 2005. Em diferentes conversas, deputados à época apareciam supostamente negociando propina com o governador. Leia mais (17/11/2008 - 21h23)

Jovem indiano morre engasgado em concurso para comer bolos

Um jovem indiano morreu engasgado durante um concurso no qual os participantes tinham que comer a maior quantidade de bolo possível. Os pais do concorrente denunciaram a falta de pessoal médico no evento. "Este tipo de concurso não deveria ser realizado. Se o concurso acontece, pelo menos deveria ser sob a estrita supervisão de especialis-

tas médicas, para que a história não se repita", disse o pai da vítima, R. K. Sabharwal, citado pela agência Ians. O jovem Sorab, 22, engenheiro da companhia de telefones celulares Nokia Siemens Networks, morreu ontem quando participava do concurso em uma cantina da localidade de Gurgaon, próxima a Nova Délhi. Leia mais (20/11/2008 - 11h29)

Galeristas e colecionadores negam crise no Brasil

Se o crash assusta lá fora, o tombo de- verá ser menor aqui dentro. Pelo menos é o que esperam galeristas e colecionadores ouvidos pela Folha. Para eles, a arte latino-americana não teve o inchaço registrado no resto do mundo, mantendo seus preços no mesmo patamar ao longo dos anos. "Nós não temos bolha, não temos nenhum Damien Hirst, nenhum Jeff Koons", compara a galerista

Luisa Strina. "É uma vantagem porque não vamos nunca ter de baixar os preços de ninguém." Mas isso não se estende a todos. O galerista Jones Bergamin, diretor da Bolsa de Arte do Rio, acredita que apenas obras mais raras devem manter o preço. "Agora, trabalhos de qualidade média e de arte contemporânea vão ter que dar uma recu- nada." Leia mais (21/11/2008 - 08h28)

Sarkozy: detenção de "Txeroki" prova colaboração

O presidente francês, Nicolas Sarkozy, assegurou hoje que a detenção do suposto chefe do aparato militar da ETA, Mikel Garikoitz Aspiazu, conhecido como "Txeroki", prova "a excelente colaboração entre França e Espanha na luta contra o terrorismo basco".

Confirmada identidade de membros do ETA presos

A Guarda Civil espanhola confirmou oficialmente que os dois detidos nesta segunda na França são Garikoitz Aspiazu, conhecido como "Txeroki" e apontado como chefe dos comandos militares do grupo terrorista basco ETA, e a acusada de integrar a organização Leire López Zurutuza.

Equador: US\$ 1 mi é queimado em roubo a carro forte

Cerca de US\$ 1 milhão foi queimado no interior de um carro forte que pegou fogo nesta quinta-feira em uma estrada dos Andes equatorianos, de-

pois que ladrões colocaram dinamite perto do tanque de gasolina após não conseguirem abrir o veículo, informou a imprensa local.

Premiê britânico apóia incentivos fiscais para ativar economia

O primeiro ministro do Reino Unido, Gordon Brown, afirmou hoje que o Reino Unido está disposto a apoiar os

incentivos fiscais para ativar a economia, como pediu o Fundo Monetário Internacional (FMI) na cúpula do G20.

Figura 34: Caso de teste número 13.



News Document Example



Perto do retorno, Ronaldo diz que está com 70% de sua forma física

Em fase final de recuperação de uma cirurgia no joelho esquerdo, o atacante Ronaldo afirmou em entrevista que será exibida nesta terça-feira pelo Canal +, da Espanha, que está com 70% de sua forma física e próximo do retorno aos gramados. "Estou com 70%. Quero dizer, muito perto de voltar a jogar. [Minha recuperação] é culpa da paciência do Bruno [Mazzioti, fisioterapeuta do

jogador]. Ele chegou a me tratar até enquanto eu dormia", revelou o atacante, em trecho antecipado pela edição desta segunda do diário "As". Ronaldo rompeu o tendão patelar do joelho em fevereiro, durante o empate por 1 a 1 do Milan com o Livorno pelo Campeonato Italiano, o que o obrigou a passar por uma cirurgia em Paris, na França. Leia mais (17/11/2008 - 17h30)

Roteiro de 24h em Lisboa tem pastel, teleférico, bonde e galeria de arte; leia trecho de guia

Ter pouco tempo para conhecer uma cidade como Lisboa pode ser angustiante. Há muito o que fazer e conhecer na cidade. Diante de tantas opções, vale definir um roteiro e aproveitá-lo ao máximo, mesmo sabendo que qualquer escolha significará deixar de lado outros roteiros igualmente bons. Reprodução Guia seleciona o melhor do design, da arquitetura e do estilo em Lisboa. O "Guia Wallpaper Lisboa" pode ser angustiante. Há muito o que fazer e conhecer na cidade. Diante de tantas opções, vale definir um roteiro e aproveitá-lo ao máximo, mesmo sabendo que qualquer escolha significará deixar de lado outros roteiros igualmente bons. Reprodução Guia seleciona o melhor do design, da

arquitetura e do estilo em Lisboa. O "Guia Wallpaper Lisboa" pode ser angustiante. Há muito o que fazer e conhecer na cidade. Diante de tantas opções, vale definir um roteiro e aproveitá-lo ao máximo, mesmo sabendo que qualquer escolha significará deixar de lado outros roteiros igualmente bons. Reprodução Guia seleciona o melhor do design, da

PSOL faz ato de apoio a Protógenes e De Sanctis em Brasília

O PSOL realiza nesta quarta-feira, a partir das 12h30, uma manifestação em apoio ao delegado federal Protógenes Queiroz, que coordenou a Operação Satiagraha e depois foi afastado do caso pela cúpula da Polícia Federal, e ao juiz Fausto De Sanctis, da 6ª Vara Criminal Federal. O ato ocorrerá em frente ao STF (Supremo Tribunal Federal), no dia em que encerra-se o prazo de defesa do banqueiro Daniel Dantas na Justiça. "Momento em que nos

uniremos perante o STF para pedir a sua condenação e para referendar o trabalho executado pelo delegado Protógenes e pelo juiz De Sanctis", informa o PSOL. De Sanctis foi responsável por decretar as duas prisões temporárias de Dantas durante a Operação Satiagraha. O banqueiro foi solto, nas duas ocasiões, após habeas corpus concedidos pelo presidente do STF (Supremo Tribunal Federal), Gilmar Mendes. Leia mais (18/11/2008 - 11h32)

Preço do petróleo da Opep recua e barril é vendido a US\$ 44,05

O barril de petróleo de referência da Opep (Organização dos Países Exportadores de Petróleo) foi vendido ontem por US\$ 44,05, após cair US\$ 1,83 em relação à cotação anterior, informou nesta sexta-feira o secretariado da organização em Viena (Áustria). Dessa forma, e com novas notícias negativas para o setor automobilístico veiculadas ontem pela imprensa, a cotação do barril da

Opep caiu para o nível mais baixo desde fevereiro de 2005. A queda de preço desde 2 de julho deste ano, quando bateu o recorde histórico de US\$ 140,73, é de 68,7%, e coincide com a sustentada tendência de queda de outros tipos de petróleo, como o Brent e o barril do petróleo leve, de referência para a Europa e para Estados Unidos, respectivamente. Leia mais (21/11/2008 - 10h00)

Wallpaper*
City Guide



Conheça o melhor da arquitetura, design, compras e vida noturna em Lisboa



Dudu Nobre e Adriana Bombom acusam companhia aérea de racismo e agressão

Adriana Bombom e Dudu nobre acusam companhia aérea de racismo

A modelo e dançarina Adriana Bombom e seu marido, o sambista Dudu Nobre, acusam a companhia aérea American Airlines de agressão e racismo. Bombom afirmou à Folha Online que os problemas começaram ainda nos Estados Unidos, quando ela tentou abrir a porta do banheiro com o avião em terra. Como o com-

partimento estava travado, uma comissária se aproximou dela e a ofendeu. "Ela me chamou de estúpida, de ignorante, de tudo", afirmou a dançarina. Renato Rocha Miranda/TV Globo Dudu Nobre e Adriana Bombom acusam companhia aérea de racismo e agressão Leia mais (18/11/2008 - 18h01)

Blog do Josias: Governo prevê PIB e arrecadação menores para 2009

Em meio à recessão das maiores economias do mundo e a crise financeira global, o governo alterou as previsões do PIB (Produto Interno Bruto) e da arrecadação para 2009, informa o blog do Josias. O crescimento do PIB no próximo ano, estimado inicialmente em 4,5%, caiu para 4%, o que implica em menos R\$ 15 bilhões de receitas e gastos públicos para o ano que vem. A perspec-

tiva do IPCA (Índice de Preços ao Consumidor Amplo) de 4,5% de inflação para 2009 subiu para 5,19%. Da queda na arrecadação, R\$ 10 bilhões, referem-se à diminuição na cotaleta de impostos federais e R\$ 5 bilhões devem-se aos royalties de petróleo. Os novos dados foram remetidos ao Congresso pelo ministro Paulo Bernardo (Planejamento). Leia mais (20/11/2008 - 03h29)

Itália planeja 1ª conferência sobre crime organizado

O governo italiano, liderado pelo primeiro ministro Silvio Berlusconi, planeja organizar a primeira conferência internacional sobre crime organizado, centrada no tráfico de dro-

gas, para levar propostas concretas ao Grupo dos Oito (G8, os sete países mais ricos e a Rússia), já que a Itália assume a Presidência do fórum a partir de janeiro.

Maconha pode preservar memória na velhice, diz estudo

Cientistas americanos dizem que substâncias presentes na maconha podem ser benéficas para o cérebro à medida que as pessoas envelhecem, reduzindo índices de inflamação e estimulando a formação de novos neurônios.

Afeganistão: combates matam 6 soldados e 5 talibãs

Um violento combate registrado na região de Farah, no oeste do Afeganistão, deixou pelo menos seis soldados afegãos e cinco supostos talibãs mortos, informou a agência EFE de uma fonte oficial.

Figura 35: Caso de teste número 14.



News Document Example



Após compra da Nossa Caixa, BB busca aquisições agora no setor privado

O Banco do Brasil não deve recuperar tão cedo a liderança no ranking dos maiores bancos no país nem com uma possível compra de 49% do banco Votorantim, braço financeiro do grupo industrial da família Ermírio de Moraes. Isso porque as atuais regras contábeis impediriam que o BB consolide em seu balanço os ativos do Votorantim, caso fique mesmo como acionista minoritário. Sétima maior instituição do país, o banco Votorantim tem o tamanho certo para devolver a liderança ao BB. Em junho, o banco somava R\$ 73,443 bilhões

em ativos, que, somados aos R\$ 512,4 bilhões de BB e Nossa Caixa, já chegariam a R\$ 585,84 bilhões – mais do que suficientes para passar os R\$ 575 bilhões de Itaú e Unibanco. Segundo o professor Ariovaldo dos Santos, da Fipecafi (Fundação Instituto de Pesquisas Contábeis, Atuariais e Financeiras), a única forma de o BB trazer todos os ativos – incluindo crédito e posição de títulos em tesouraria – para o seu balanço é formalizar um acordo em que tenha a gestão compartilhada do novo banco. Leia mais (21/11/2008 - 09h53)

Diretor-geral da PF descarta possibilidade de Protógenes reassumir caso contra Dantas

O diretor-geral da Polícia Federal (PF), Luiz Fernando Corrêa, afastou nesta terça-feira a possibilidade do delegado Protógenes Queiroz reassumir a coordenação das investigações relativas à Operação Satiagraha. Corrêa disse ainda que não são “personalizadas” as ações da PF e que Protógenes é “um servidor a mais na Polícia Federal”. “A investigação [da Satiagraha] já foi redirecionada e tem um novo delegado e uma nova equipe interagindo com o Mi-

nistério Público e com a Justiça produzindo diligências”, afirmou Corrêa após cerimônia no Ministério da Justiça, afastando a hipótese de Protógenes retomar a coordenação da operação. Ao ser questionado sobre o futuro de Protógenes, o diretor foi evasivo. “Ele [Protógenes] é um profissional de polícia. Ele é um policial, enquanto ele estiver na PF. Protógenes é um servidor a mais na PF”, afirmou Corrêa. Leia mais (18/11/2008 - 11h21)

Defesa Civil permanece em atenção devido a risco de mais chuva no Rio

Depois da chuva que alagou vias e prejudicou o trânsito no Rio ontem, a cidade amanheceu nesta terça-feira com sol e poucas nuvens. No entanto, ainda há previsão de chuvas rápidas, e a Defesa Civil permanece em estado de atenção. No fim da tarde de segunda-feira (17), uma chuva forte formou pontos de alagamento em vias de toda a cidade, o que gerou congestionamento. O aeroporto Santos Dumont (centro do Rio) fechou por cerca de uma hora, e a Ponte Rio-Niterói, por 20 minutos. Nesta terça, os aeroportos Santos Dumont e internacional Tom Jobim, na Ilha do Governador (zona norte), operam normalmente, segundo a Infraero. O trânsito na cidade também tem poucas retenções, de acordo com a CET-Rio. Leia mais (18/11/2008 - 12h12)

O aeroporto Santos Dumont (centro do Rio) fechou por cerca de uma hora, e a Ponte Rio-Niterói, por 20 minutos. Nesta terça, os aeroportos Santos Dumont e internacional Tom Jobim, na Ilha do Governador (zona norte), operam normalmente, segundo a Infraero. O trânsito na cidade também tem poucas retenções, de acordo com a CET-Rio. Leia mais (18/11/2008 - 12h12)

Paulistanos podem assistir a 12 obras do cineasta alemão Murnau

Divulgação Friedrich Wilhelm Murnau (foto) iria completar 120 anos em 2008. O público paulistano tem até o dia 30 de novembro para assistir aos 12 filmes produzidos pelo cineasta Friedrich Wilhelm Murnau (1888-1931), considerado um dos mais importantes nomes do cinema mudo e do expres-

sionismo alemão. O Centro Cultural Banco do Brasil, na região central da cidade de São Paulo, organizou uma programação que promove uma retrospectiva completa do trabalho do artista, que completaria seu 120º aniversário neste ano. Leia mais (20/11/2008 - 07h42)



Friedrich Wilhelm Murnau (foto) iria completar seu 120º aniversário neste ano

Itália planeja 1ª conferência sobre crime organizado

O governo italiano, liderado pelo primeiro ministro Silvio Berlusconi, planeja organizar a primeira conferência internacional sobre crime organizado, centrada no tráfico de drogas, para levar propostas concretas ao Grupo dos Oito (G8, os sete países mais ricos e a Rússia), já que a Itália assume a Presidência do fórum a partir de janeiro.

Presidente do Santos admite vender Kléber Pereira após o Nacional

O presidente do Santos, Marcelo Teixeira, admitiu que poderá vender o atacante Kléber Pereira, um dos principais jogadores da equipe nesta temporada, depois do Campeonato Brasileiro. No entanto, ele ressaltou que ainda conversará com o atleta. “Temos que esperar o campeonato acabar e saber qual é o desejo do jogador. Não queremos prejudicar a carreira dele.

Caso ele queira sair, a proposta precisa ser boa para o clube também”, disse o dirigente. Daniel Pereira, irmão e procurador do atacante, afirmou que o centroavante recebeu uma boa proposta para se transferir para um clube do exterior. Mas, Marcelo Teixeira destacou que o Santos ainda tem contrato com o atleta. Leia mais (20/11/2008 - 11h53)

Colar de 2 mil anos é avaliado em R\$ 1,2 milhões

Uma gargantilha de ouro de cerca de 2 mil anos avaliada em 350 mil libras esterlinas (cerca de R\$ 1,2 milhões) é en-

contrada por um amador está sendo considerada um dos mais importantes achados do tipo em 50 anos na Grã Bretanha.

Cientistas observam como buraco negro da Via Láctea devora matéria

Uma equipe internacional de cientistas observou como o buraco negro supermassivo da Via Láctea, com uma massa 4 milhões de vezes maior que a do Sol, segue devorando matéria no centro da galáxia. A descoberta foi anunciada pelo Centro Superior de Pesquisas Científicas da Espanha (Csic). Os pesquisadores detectaram intensos brilhos de Sagittarius A (SgrA), o bu-

racos negro supermassivo, situado no centro da galáxia, a 26 mil anos-luz da Terra, que revelariam a existência de nuvens de gás desgarradas ao girar em grande velocidade ao redor da formação espacial. Divulgação Concepção artística mostra nuvens de gás em torno de buraco negro; formação continua devorando matéria no centro da galáxia. Leia mais (18/11/2008 - 19h00)

AGU vai à Justiça para que Abin acompanhe perícia da PF

O advogado-geral da União, José Antonio Dias Toffoli, vai pedir ao juiz Ali Mazloum, da 7ª Vara Federal Criminal, que reconsidere o despacho em que proíbe que oficiais da Abin (Agência Brasileira de Inteligência) acompanhem com o Ministério Público a abertura do laço e a verificação de documentos que foram apreendidos na sede

da agência pela Polícia Federal. O juiz proibiu também o acesso da Abin aos exames técnicos em celulares e notebooks recolhidos com arapongas e com o delegado Protógenes Queiroz. Os documentos que estão em posse da PF são considerados secretos e originalmente destinados ao presidente da República. Leia mais (17/11/2008 - 13h24)



Concepção artística mostra nuvens de gás em torno de buraco negro; pesquisadores afirmam que formação continua devorando matéria no centro da galáxia

Figura 36: Caso de teste número 15.



News Document Example



Polícia investiga hipótese de suicídio em caso de mulher que caiu de prédio em Guarulhos (SP)

A Polícia Civil afirma que trabalha também com a possibilidade de que a operadora de caixa Andréia Cristina Nóbrega, 31, possa ter cometido suicídio após uma discussão com o ex-marido. Ela morreu após cair da janela do terceiro andar do prédio onde morava, em Guarulhos (Grande São Paulo), na terça-feira (18). O filho dela, de 6 anos, caiu na marquise e está internado. Ex-marido de Andréia e pai

Shiloh quer ser chamada de John, diz Brad Pitt

Em entrevista a Oprah Winfrey, Brad Pitt revelou uma mania da filha de dois anos e meio, Shiloh. Lionel Cironneau/AP Brad Pitt com Zahara (à dir.) e Shiloh; "An-



Com os filhos Zahara (à dir.) e Shiloh, Brad Pitt chega ao hospital onde está Angelina

Feriado tem programações especiais nos museus de SP

O município de São Paulo aderiu nesta quinta-feira ao feriado pelo Dia da Consciência Negra. Atracões especiais acontecem por toda a cidade para a comemoração. O Museu do Futebol programou uma visita especial para o feriado com foco na história de aceitação dos negros, mulatos e mestiços no esporte. No Museu Afro Brasil, duas novas exposições têm início neste 20 de novembro, além de

a apresentação de vídeos e performances culturais para comemorar a data. Outra opção é a apresentação de dança prevista para acontecer no Museu da Casa Brasileira, que deve contar com a presença do grupo perucial "Chocolate Riveros y Los Canjones". Também devem acontecer exposições de filmes, como o do documentário "Philadelphia - Orixá: Yomonjá", que aborda o racismo. Leia mais (20/11/2008 - 11h46)

Inflação em SP fica 0,58% na segunda prévia de novembro, diz Fipe

O IPC (Índice de Preços ao Consumidor) no município de São Paulo registrou uma alta de 0,58% na segunda quadrissemana de novembro -30 dias até 15/11-, ligeira correção para cima em relação ao índice visto na abertura do mês, 0,57%. Os dados foram divulgados nesta terça-feira pela Fipe (Fundação Instituto de Pesquisas Econômicas). En-

tenda a diferença entre os principais índices de inflação A categoria Alimentação registrou a maior variação percentual de preços, 70%, mas a contribuição para a composição final do indicador foi de 27,28%. Já a categoria Habitação teve uma variação ligeiramente menor nos preços, 0,69%, mas teve maior peso na composição do índice, 39,08%. Leia mais (18/11/2008 - 08h46)

Cientistas observam buraco negro da Via Láctea

Uma equipe internacional de cientistas observou como o buraco negro supermassivo da Via Láctea, com uma massa 4 milhões de vezes maior que a do

Sol, segue devorando matéria no centro da galáxia, informou nesta terça o Centro Superior de Pesquisas Científicas da Espanha (Csic).

PUBLIFOLHA: Livro explica a resistência à reforma tributária e a alta de impostos no Brasil

"Em menos de dez anos a carga tributária deixou o patamar de 25% da renda nacional, no qual havia se mantido por mais de duas décadas, e se aproxima da marca de 36%." A mesma do Reino Unido - Gustavo Patu, em "A Escalada da Carga Tributária" Divulgação Gustavo Patu explica como a carga tributária cres-

ceu tanto no Brasil Como é possível que, em menos de uma década, durante tempos de paz, democracia e em plena hegemonia do pensamento liberal, a carga tributária do Brasil tenha crescido tanto? E por que motivos há tanta resistência à realização de uma reforma tributária no país? Leia mais (21/11/2008 - 09h47)

Sarcófagos achados no palácio de Herodes são exibidos

Três sarcófagos com cerca de 2 mil anos descobertos em escavações no palácio de Herodes, rei da Judéia (37 a.C. e 4 a.C.), foram exibidos na Universidade He-

Gene pode explicar dependência de cocaína

Parte da população mundial pode ser mais vulnerável ao vício de cocaína devido a uma variação genética, de acordo com uma nova pesquisa feita por cientistas de diversos países.



A Escalada da Carga Tributária Gustavo Patu

Gustavo Patu explica como a carga tributária cresceu tanto no Brasil

Figura 37: Caso de teste número 16.



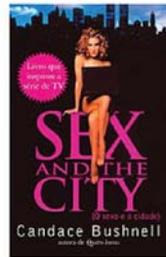
News Document Example



Bolsa de NY cai mais de 5% e fecha no menor nível em cinco anos

A Bolsa de Nova York fechou em nova queda nesta quinta-feira, que a levou ao seu nível mais baixo em cinco anos e meio, afetada pela degradação da economia norte-americana e pela ausência de uma resposta política à crise na indústria automobilística. O índice Dow Jones Industrial Average, da Nyse (Bolsa de Valores de Nova York, na sigla em inglês), recuou 5,56%, aos 7.552,29 pontos -no seu nível mais

baixo desde março de 2003. Em um pregão volátil, o Dow Jones chegou a descer 6,1% e a subir 2,37%. O índice Nasdaq, de elevado componente tecnológico, perdeu 5,07%, aos 1.316,12 pontos. Já o S&P 500 caiu 6,71%, a 752,44 pontos -nível não registrado desde abril de 1997. "É muito perigoso. O mercado está cauteloso", disse Lindsey Piegza, da FTN Financial. Leia mais (20/11/2008 - 20h26)



Livro traz as histórias picantes de Carrie, Samantha, Miranda e Charlotte

LIVRARIA: Veja como foi o encontro de Carrie e Mr. Big em "Sex and The City"

A narrativa do primeiro encontro de Carrie e Mr. Big está no livro "Sex and The City", que traz as histórias originais que deram origem à série de TV e ao longa-metragem de mesmo nome. Reprodução Livro traz as histórias originais de "Sex and The City" em 352 páginas. O livro já vendeu mais de um milhão de exemplares nos EUA e tem edição brasileira da Editora Record, disponível na Livraria da Folha. Leia mais (20/11/2008 - 20h16)

STJ veta alta em plano de saúde para idosos associados da Unimed Natal; ouça

Na semana passada, o STJ (Superior Tribunal de Justiça) proibiu o aumento das mensalidades de planos de saúde por faixa etária para quem tem 60 anos ou mais. O entendimento vale apenas para os associados da Unimed Natal. Maria Inês Dolci, coordenadora institucional da Pro Teste, colunista da Folha e responsável por um blog da

Folha Online, diz que a empresa informou aos seus clientes de planos antigos que a partir de 2004 iria aumentar em 100% e 200% as mensalidades para quem completasse 60 e 70 anos, respectivamente. Ouça outros podcasts com a participação da advogada. Leia mais (17/11/2008 - 12h00)

EUA: após incêndios, ator encontra casa destruída

Obrigado a deus, Christopher Lloyd encontrou a sua casa destruída após dois incêndios que atingiram a região de Montecito, no Estado americano da Califórnia, informou o ator Ch-

ristopher Lloyd voltou ontem e encontrou a residência de US\$ 11 milhões completamente destruída, informou a rede CNN.

Maranhão diz que espera publicação de decisão do TSE para assumir governo da Paraíba

O senador José Maranhão (PMDB-PB) disse nesta sexta-feira que aguarda a publicação do acórdão do TSE (Tribunal Superior Eleitoral) para renunciar ao cargo e assumir o governo da Paraíba. Sorridente, Maranhão afirmou que a decisão chegou "com grande retardo", mas confirma a vitória que ele disse merecer. Ele também negou que tenha sonhado informações sobre sua evolução patrimonial. "Todo meu patrimônio está

declarado e claramente demonstrado. Todos os homens públicos correm esses riscos [de supostas falsas denúncias]", afirmou Maranhão, que antes de ir ao seu gabinete se reuniu com os advogados de defesa. O senador se referiu às informações, que vieram à tona no ano passado, de que ele não informou à Receita Federal sobre sua evolução patrimonial, deixando de declarar 28.290 cabeças de gado. Leia mais (21/11/2008 - 13h42)

Posição de carrinho 'pode afetar desenvolvimento do bebê'

Bebês e crianças que são levadas em carrinhos que as deixam de costas para os pais podem ter seu desenvolvimento prejudicado, sugere um estudo da Universidade de Dundee, na Escócia.

Conselho Eleitoral marca eleições no Equador

As eleições gerais no Equador ocorrerão em 26 de abril de 2009, informou nesta terça-feira o presidente do Conselho Nacional Eleitoral (CNE), Omar Simons.

Figura 38: Caso de teste número 17.



News Document Example



Dilma diz que investir no pré-sal é questão anticíclica para Brasil

A ministra Dilma Rousseff (Casa Civil) disse nesta sexta-feira que os investimentos na camada do pré-sal têm caráter anticíclico contra a crise mundial e que manter o volume elevado de recursos na exploração dessa área permite que o Brasil aproveite as “oportunidades de crescimento” no futuro. “A descoberta de hoje, para nós, é muito importante porque configura um dos fatores anticrise no Brasil. A decisão do governo de manter os investimentos no PAC [Programa de Aceleração do Crescimento] e no pré-sal é importante porque são ques-

tões eminentemente anticíclicos”, disse ela, numa referência ao anúncio de novas reservas na camada do pré-sal no Espírito Santo, divulgado hoje pela Petrobras. Ainda comentando as descobertas de hoje, a ministra afirmou: “estamos caminhando na direção de constituir o pré-sal como uma das opções para os próximos anos, para as próximas décadas. Por isso, o governo mantém a decisão de investir no pré-sal e acha que o investimento dessa qualidade é algo fundamental como oportunidade de crescimento para o país”. Leia mais (21/11/2008 - 13h54)

Ban expressa a preocupação com Gaza

O secretário geral da ONU, Ban Ki moon, ligou nesta terça-feira ao primeiro ministro de Israel, Ehud Olmert, para expressar sua “profunda preocupação” pela situação humanitária em Gaza.

Bomba mata líder de família mafiosa em Tel Aviv

Yacov Alperon, considerado o chefe da maior família ligada ao crime organizado em Israel, foi assassinado nesta segunda-feira, quando uma bomba explodiu no carro em que estava, em uma das ruas principais de Tel Aviv.

Chip's Burger lança sanduíche para acompanhar balada em SP

Mauro Holanda/Divulgação. O Chip's Burger tem a intenção de recriar uma atmosfera norte-americana, serve em seu cardápio sanduíches clássicos, além de uma nova versão: o pernil da madrugada. Leia mais (20/11/2008 - 20h44)



Sanduíche de pernil do Chip's Burger é servido às sextas e sábados, na madrugada

Emenda pode cortar verba de partido neonazista alemão

Os ministros de Interior do governo alemão e os Estados federados pretendem cortar o financiamento estatal aos partidos nazistas de extrema direita, por meio de uma emenda na Constituição que vetaria o repasse dos fundos.

Milícia governista e rebeldes enfrentam-se no Congo um dia após início de retirada

Os rebeldes congolese liderados por Laurent Nkunda travaram combates nesta quinta-feira com milícias favoráveis ao governo na região leste da República Democrática do Congo (RDC, antigo Zaire), um dia depois de um recuo rebelde na linha de combate. “Os en-

partidos nazistas de extrema direita, por meio de uma emenda na Constituição que vetaria o repasse dos fundos. frentamentos começaram esta manhã nas cidades de Katoro e Nyongera, 5 km ao norte de Kiwanja”, afirmou uma fonte dos serviços de segurança, que pediu anonimato. Kiwanja fica 75 km ao norte de Goma, a capital da província de Kivu Norte. Leia mais (20/11/2008 - 11h19)

Figura 39: Caso de teste número 18.



News Document Example



Jobim se reúne com Lula para discutir Plano Nacional de Defesa

O governo federal prepara uma série de discussões ainda neste ano sobre o Plano Nacional de Defesa. A medida, segundo o ministro Nelson Jobim (Defesa), vai rever a obrigatoriedade do serviço militar no país, o fortalecimento da indústria bélica e a recolocação dos militares na sociedade, entre outros pontos. A agenda para o plano começa na terça-feira, quando ele apresentará a proposta a parlamentares da Câmara e do Senado. Na

quinta-feira, Jobim tem uma reunião no Palácio do Planalto com o presidente Luiz Inácio Lula da Silva e ministros Mangabeira Unger (Assuntos Estratégicos), Dilma Rousseff (Casa Civil) e Franklin Martins (Comunicação) para concluir o plano. Em seguida, Jobim convocará o Conselho de Defesa Nacional. O ministro estima aprovar o plano até 15 de dezembro. Leia mais (20/11/2008 - 18h17)

Índia afirma que G20 superou G7 na resolução de problemas econômicos

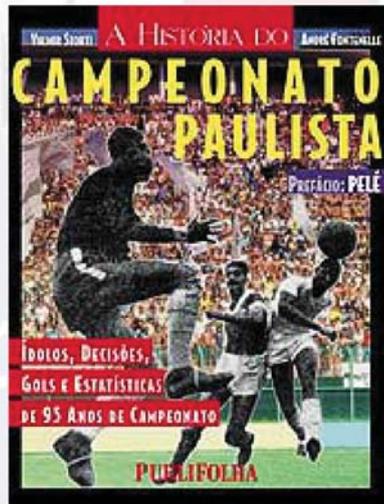
O ministro de Finanças indiano, Palaniappan Chidambaram, disse nesta terça-feira que o G20 (grupo que reúne os países mais ricos e os principais emergentes) transformou-se no fórum mundial mais importante para resolver problemas econômicos. A declaração indiana segue a linha adotada pelo presidente Luiz Inácio Lula da Silva, mas que encontra resistência por parte do primeiro-ministro italiano Silvio Berlusconi. "O G7 (os sete países mais ricos do mundo) reconheceu tardiamente que não tem a solução para todos os problemas", disse Chidambaram durante a realização do Fórum Econômico Indiano, segundo um comunicado emitido pela organização. Leia mais (18/11/2008 - 13h50)

Polícia prende dez pessoas em tentativa de assalto a transportadora em Canoas (RS)

Policiais da Brigada Militar prenderam dez pessoas na madrugada desta sexta-feira após uma tentativa de assalto à transportadora Transpaulo, no bairro de Matias Velho, em Canoas (RS). De acordo com a polícia, um grupo fortemente armado invadiu a empresa por volta das 0h30 e rendeu cerca de 60 funcionários. Eles obrigaram os empregados a carregar três caminhões com mercadorias. A polícia foi avisada e cercou a transportadora. Na tentativa de fuga houve troca de tiros e dez suspeitos foram presos, entre eles, oito que tentaram fugir em um Fiesta. Leia mais (21/11/2008 - 10h25)

AIEA confirma que há urânio em instalação síria

A Agência Internacional de Energia Atômica (AIEA) confirmou hoje em Viena que foram encontrados rastros de urânio em uma suposta instalação militar na Síria, bombardeada ano passado pela aviação israelense e onde os Estados Unidos asseguram que ia ser construído um reator nuclear.



Livro traz ídolos, decisões, Gols e estatísticas de 95 anos de Campeonato

LIVRARIA: Mais de 100 livros de R\$ 1,99 a R\$ 9,90; há viagens, filhos, carreira e receitas

Veja abaixo a lista completa de mais de 100 títulos em promoção na Livraria da Folha, uma lista de R\$ 1,99 a R\$ 9,90. A lista é atualizada diariamente com novas promoções. Confira também

Figura 40: Caso de teste número 19.



News Document Example



Novela: Luca convida Dani para trabalhar no Nonna Mia em “Água na Boca”

A programação está sujeita a alteração devido à edição dos capítulos e é de responsabilidade da emissora Água na Boca (Band, 20h10) Keiko (Eda Nagayama) decide enfrentar Artur (Peterlongo/João Signorelli) e conta com a ajuda de Kim (Alexandre Moreno). Françoise (Jacqueline Laurence) e Jean-Paul (Mário César Camargo) chegam ao Mercado. Zeca (Joaquim Lopes) encara Françoise. Miquelina (Ana Cecília) diz para Severino (Claudio Jaborandy) que irá filmar o 'chamego' com Bertinho (Betito Tavares). Luca (Caetano O'Maihlán) pede para as Mamas cuidarem de Allegra (Karen Tavares). Jean-Paul conta para Zeca que os Cassou-

let perderam tudo. Érika (Marisol Ribeiro) conta para Alex que Luca tem uma filha. Ele quer saber se ela está pronta para trabalhar no Paris. Dani sai em busca de emprego, mas ninguém a aceita. Alex propõe a Remi (Marco Antônio Pâmio) que volte a trabalhar no Paris. Dani descobre que Alex divulgou o escândalo de sua família no jornal. Alex convoca uma reunião com os funcionários do Paris. Luca vai atrás de Gina (Miranda Kassin) para saber porque ela escondeu Allegra por tantos anos. Bertinho chega ao Baião de Dois. Alex quer recontratar os funcionários do Paris e convida Érika para ser chef. Keiko mostra a foto de Artur para Kim. Severino vai atrás de Miquelina e Berti-

nho, que estão no quarto. Os dois fingem que estão transando. Érika finge surpresa a proposta de Alex. Dani entra no Bistrô e parte para cima dele. Severino (Claudio Jaborandy) vê Miquelina (Ana Cecília) e Bertinho (Betito Tavares) na cama. Ele sai magoado e Miquelina se arrende. Dani (Rosanne Mulholland) diz para Alex (Alexandre Barros) que não vai desistir de sua vida. Érika (Marisol Ribeiro) leva Dani para a sua casa. Luca (Caetano O'Maihlán) convida Dani para trabalhar no Nonna Mia. Ela promete pensar. Allegra (Karen Tavares) tem um pesadelo e vai dormir com Luca e Érika, que fica uma fera. Leia mais (18/11/2008 - 00h16)



"Em reunião com Barack Obama, McCain prometeu 'unir forças' para reformas" <http://www1.folha.uol.com.br/folha/mundo/ult9pg06628.shtml>

Após derrota nas eleições, McCain volta ao Senado

O ex-candidato republicano à Casa Branca, John McCain, senador do Arizona, voltou nesta terça-feira, 12.nov.08/Efe Em reunião com Barack Obama, McCain prometeu "unir forças" para reformas O líder da minoria republicana, o senador Mitch McConnell, afirmou estar "contente" com o retorno do ex-candidato à Presidência. Leia mais (18/11/2008 - 18h00)

Ruas e avenidas do Rio apresentam tráfego intenso nesta terça-feira

As principais ruas e avenidas no Rio apresentam tráfego intenso na tarde desta terça-feira, segundo a prefeitura. Por volta das 17h30, havia lentidão no centro e nas regiões norte e sul da cidade. O motorista reduzia a velocidade no centro, na avenida Perimetral, e na altura da Praça Mauá, como na altura da

rua Professor Pereira Reis, ambos os trechos no sentido avenida Brasil; e na rua Frei Caneca, altura da rua do Santana. Na zona norte, o acesso às vias estava complicado no entorno da praça da Bandeira, no sentido Tijuca; e na avenida Brasil, altura da rua Prefeito Olímpio de Melo, na pista lateral, no sentido zona oeste Leia mais (18/11/2008 - 17h40)

PUBLIFOLHA: Veja obras da série "Folha Explica" em promoção; títulos a partir de R\$ 3,90

Quem gosta de ler deve aproveitar a promoção da Publifolha. Veja a lista de livros da série "Folha Explica" em oferta por tempo limitado a partir de R\$ 3,90 e abasteça a sua biblioteca. O preço normal dos livros varia entre R\$ 17,90 e R\$ 20,90. Procure livros de outros temas AQUI. * Leia mais (18/11/2008 - 13h11)

ONU: Rice participa de combate a violência contra mulher

A secretária de Estado americana, Condoleezza Rice, assinou uma campanha promovida pela ONU para pressionar os governos domundo a acabar com a violência contra as mulheres.

Figura 41: Caso de teste número 20.

Apêndice B Formulários para a avaliação qualitativa

Neste apêndice encontram-se os modelos de formulários, que são a carta de introdução à avaliação empírica (Figura 42), o questionário pré-teste (Figura 43) e a grade de resposta (Figura 44), utilizados para a avaliação empírica deste trabalho, como descrito no Capítulo 5.

CPPH – Centro de Pesquisa PUCRS/HP
 Faculdade de Informática/PUCRS
 Avenida Ipiranga, 6681 – Prédio 32, sala 101 – Porto Alegre – RS
 Tel: (51) 3320-3695
 Pesquisadores Responsáveis:
 Prof. Dr. João Batista Souza de Oliveira e Ricardo Farias Piccoli
joao.souza@pucrs.br, rfbpiccoli@gmail.com

Ao avaliador

Antes de tudo, muito obrigado por participar deste teste. Este trabalho está sendo desenvolvido em uma parceria entre FAMECOS/FACIN/HP, e em parte como minha dissertação de mestrado pela Faculdade de Informática.

O objetivo é investigar a viabilidade de se fazer diagramação automática de documentos, neste caso documentos similares a páginas de jornal. Um exemplo de utilização seriam *newsletters* diários em empresas, folhetos e guias diários, etc, ou seja, qualquer aplicação em que não haja tempo suficiente para um *designer* humano construir páginas, frente ao volume de conteúdo produzido em um tempo muito curto. Um *layout* de jornal é um bom modelo para se trabalhar com produção automática, porque o formato da página é simples e familiar para o leitor, o que não lhe causa surpresas ou dificuldades na hora de ler o conteúdo.

Como parte da validação deste trabalho, é necessária uma avaliação experimental dos resultados gerados por pessoas – neste caso, você – que tenham experiência ou conhecimento em projeto gráfico de páginas.

A avaliação é bastante simples: você tem 20 exemplos de páginas de jornal construídas automaticamente. Cada uma foi montada com notícias de um serviço de notícias pela *internet* (RSSs da Folha de São Paulo e do Terra). As notícias e figuras contidas nas páginas não foram planejadas para serem impressas em uma folha de jornal, logo, é normal que muitas dessas notícias possuam títulos compridos demais ou textos muito curtos. O único critério de avaliação é o aspecto geral do documento. Para cada uma das páginas, observe-a e tente responder a pergunta a seguir:

Em uma escala de 1 a 5 (sendo 5 a melhor nota), que nota você daria para o aspecto geral desta página, considerando sua aparência e estética?

Caso você tenha algum comentário a respeito da qualidade de cada página, sinta-se livre para expressar-se. Note que devem ser considerados aspectos relacionados apenas à **forma** de cada página, e não ao **conteúdo** que está inserido (tamanho dos títulos, texto, número de parágrafos, legendas nas fotos, etc.), visto que o sistema não tem controle sobre as amostras geradas.

Muito obrigado pela sua participação!

Os dados aqui informados serão utilizados para a finalidade desta pesquisa e como base para futuras publicações e divulgações do tema.

O anonimato dos entrevistados será preservado em todo e qualquer documento divulgado em foros científicos (tais como conferências, periódicos, livros e assemelhados) ou pedagógicos (tais como apostilas de cursos, slides de apresentações e assemelhados).

Por favor, indique sua posição em relação aos termos acima:

Estou de pleno acordo com os termos acima.

Assinatura: _____

Figura 42: Formulário para participação dos voluntários na avaliação empírica deste trabalho (Capítulo 5).

Pontifícia Universidade Católica do Rio Grande do Sul
Faculdade de Informática
Centro de Pesquisa PUCRS/HP – Laboratório APL
Responsável: Professor Dr. João Batista Souza de Oliveira

Diagramação Automática de Documentos

Questionário Pré-Teste

Data: __/__/__

Nome (opcional): _____ Idade: _____ anos

Sexo: () Masculino () Feminino

Profissão: _____

1. Com que frequência você lê jornais (mídia impressa)?
() Uma vez por semana
() Duas vezes por semana
() Três vezes por semana
() Mais de três vezes por semana
2. Quantas publicações diferentes de jornais você costuma ler?
() Um jornal
() De 2 a 3 jornais diferentes
() Mais de 3 jornais diferentes
3. Há quantos semestres você está na faculdade? _____
4. Quantas disciplinas relacionadas a *design* gráfico você já cursou? _____
5. Você já trabalha ou trabalhou com *design* gráfico? Se sim, há quanto tempo?
() Nunca trabalhei com *design* gráfico
() Até 1 ano de experiência
() De 1 a 2 anos de experiência
() Mais de 2 anos de experiência

Figura 43: Questionário pré-teste para avaliação empírica (Capítulo 5).

Para cada página, marque um 'X' na coluna correspondente à sua avaliação:

		Qualidade da página				
		1 (Muito Ruim)	2 (Ruim)	3 (Razoável)	4 (Bom)	5 (Muito Bom)
Número da página	1					
	2					
	3					
	4					
	5					
	6					
	7					
	8					
	9					
	10					
	11					
	12					
	13					
	14					
	15					
	16					
	17					
	18					
	19					
	20					

Caso deseje fazer algum comentário a respeito dos resultados, utilize o espaço abaixo, ou o verso da folha:

Figura 44: Tabela para avaliação empírica dos casos de teste (Capítulo 5).