

HTMLビューワの表示に関する一考察

—HTML文書の構造化と表示—

5N-1

村本 貴英 鈴木 利光 齊藤 一実

(株)富士通研究所

1. はじめに

WWWビューワであるMosaicは文字や絵をマウスでクリックするだけで、世界中の公開情報に簡単にアクセスでき、それまでインターネットに関心がなかった人々を引き付けた。しかし、情報記述言語(HTML)の仕様をそのまま再現するだけでは情報のレイアウト能力に限界があり、また動画や静止画を表示するにはWWWビューワとは別にアプリケーションを用意する必要がある。今後、本格的な商用サービスを大衆向けに提供し発展させていくには、より柔軟かつ効率良く情報をレイアウト/表示でき、サービスに合わせてGUIをカスタマイズできるようなビューワが必要であろう。

これらの問題に対して、HTMLの仕様を拡張せずにWWWクライアント側で描画レイアウト機能を向上させることを目的とし、HTML文書を階層構造化して管理することで各種情報の柔軟なレイアウト表示を可能にするWWWクライアント「Brick」を開発した。また各種メディアデコーダを内蔵することによりマルチメディア情報の一元表示を可能にした。

2. HTML文書の階層構造化の管理

HTML文書はその文書構造を複数レベルのヘッダエレメントで章だてや節を定義してある(これを文書ブロックと呼ぶ)。Brickでは、まず文書ファイルデータを文書ブロックごとに分割し、更に各文書ブロック中の文章のかたまり(これを文書パラグラフと呼ぶ)ごとに分割する(図1参照)。そして、上位

A Study of HTML Viewer
:Structuring and Viewing of HTML Documents
Takahide Muramoto, Toshimitsu Suzuki,
Kazumi Saito
Fujitsu Laboratories Ltd.
64, Nishiwaki, Ohkubo-cho, Akashi 674, Japan

の階層の文書ブロックから順に文書パラグラフを描画していく。これにより、下位ヘッダ情報のボタン化やイメージの側面への文字の流し込みが可能になり、限られた表示領域を有効利用できる。このように文書の階層構造と表示領域を管理することでより柔軟なレイアウトを可能にした。

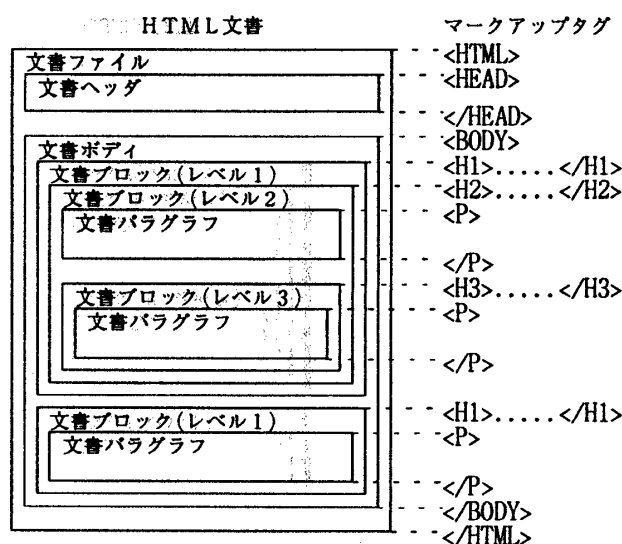


図1: HTML文書の階層構造

3. 内部構成

Brick内部のモジュール構成を以下に示す(図2参照)。

- (1)通信部: HTTPによるデータ転送を行なう。
- (2)キャッシュコントローラ: 通信部を介してネットワーク上から取得されたデータをローカルファイルとして保存し、データのヘッダ情報などをキャッシングして以降のネットワークアクセスを軽減する。
- (3)SGMLパーザ: 指定したDTDに従ってHTML文書のマークアップを解析し、エラーチェックを行なった上でエレメントの抽出を行なう。
- (4)文書要素管理部: SGMLパーザでデコードした結果をもとにHTML文書を階層構造化する。また個々の

階層関係、文書パラグラフ内の各エレメントの機能に関連した情報を記述したツリー構造の管理テーブルを作成する。

(5)レイアウト：画面上の表示領域を管理する。画面を適切なサイズ・形状の領域に分割し、レンダラを用いて文字列や画像などを描画させる。レイアウト機能モジュールの変更により論文形式の文書表示やカタログ雑誌形式など様々なレイアウトを実現できる。また個々のイベント発生に応じた処理の命令を各モジュールに伝達する。

(6)レンダラ：レイアウトから指定された領域に描画する。領域の幅に合わせて行をラッピングしたり、画像の挿入やインデント下げなどの処理をする。

(7)イベントハンドラ：ページ送りボタンやページ戻しボタン、及びアンカーや入力フィールドなどへのイベント管理をする。

4. 表示画面

実際にWWWサーバー(<http://www.fujitsu.co.jp/>)にアクセスした時のBrickの表示画面を図3に示す。

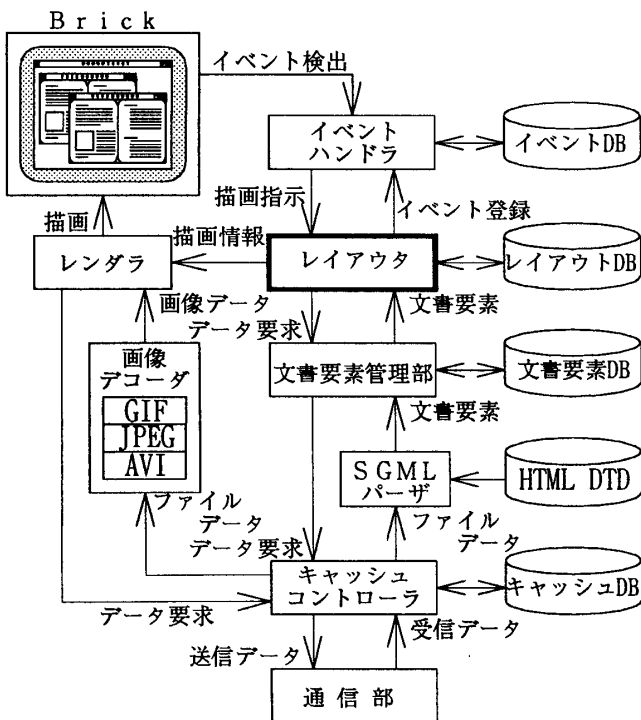


図2：Brickの内部構成図

画面の空き領域を有効利用するため見開き表示を行なっている。コンピュータに馴染みの薄い利用者でも簡単に利用できるように、ボタンによるページ切り替えで情報にアクセスできる。また、Mosaicでは外部アプリケーションに頼っていた動画像や音声データの再生を内部機能として用意し、あたかも本を眺めるように各種マルチメディアデータを統合化して扱うことが可能である。

5. おわりに

BrickではHTMLの仕様を変更せずに情報のレイアウト能力を向上させることができ、マルチメディア情報を一元的に表示可能とした。更に静的ではあるがレイアウトモジュールを変更することで情報表示レイアウトの変更を可能にしている。しかし、既存のHTML文書はMosaicをレファレンスとして記述されていることもあり、うまくレイアウトできないものもあった。またMosaicに比べてレイアウト部分の処理が余分に掛かるため描画が完了するまでにやや時間がかかる問題がある。

今後は、レイアウト機能を洗練させると共に、サービス毎に動的にレイアウトモジュールを変更できる機構の検討を行なっていく予定である。

【参考文献】

- ・HTML+ Discussion Document INTERNET-DRAFT / November 2, 1993
- ・Hypertext Transfer Protocol INTERNET-DRAFT / Tim Berners-Lee, CERN / May 5, 1994

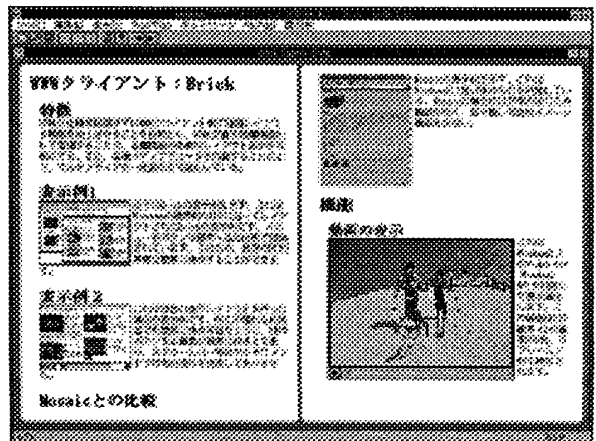


図3：Brickの画面表示