1 はじめに

最近では，新聞，TV，雑誌など様々なメディアにおいて “Internet” というキーワードを見かけることが多くなってきた．Internet とは全米科学財団 (NSF) が出資して設立された NSFNET と，それに接続されたネットワークの集合体を総称する語句である．本来は学術目的のネットワークとして発展してきたが，最近は商用ネットワークも多数接続されるようになり，ユーザ数は急激に増加している．具体的な数字を挙げると，1994 年 7 月の時点で NSFNET に接続されているホスト数は約 320 万である． ユーザ数を単純にそれの 10 倍と見積もるならば，およそ 3200 万人のユーザが何らかの形で Internet を利用しているという計算になる [1].

Internet で利用できるサービスにはいろいろあるが，主に利用されているのは電子メール，ニュース，リモートログイン，ファイル転送などである．1994 年 7 月の時点で，NSFNET バックボーンを通過するパケットの半数以上がこれらのサービスに属するものである [2].

現在では，Internet は情報を蓄積・流通させるための巨大な環境ということができるだろう．Internet 全体をデータベースサーバとして見るならば，そこに蓄積されている情報量は膨大なものである．その情報をハイパーテキストとして閲覧するという試みが World Wide Web (以降 WWW) である．本稿では WWW の概要を簡単に説明するとともに，他のメディアとの連携など，その応用について紹介する．

2 WWW の特徴

Internet についてによく話題にされる問題点の一つに，ディレクトリ情報と言えるものが整備されていないということがある．どこにどんな情報があり，どんなサービスが行われているのかを知るための情報が今一つ捉えにくいのである．また，Internet は学術目的のネットワークとして発展してきたこともあり，使いこなすにはコンピュータに関するある程度の知識が必要とされる．それに，初心者には馴染みにくいキャラクタベースのユーザインターフェースも多く使われている．従って，必ずしも全ての人にとって使いやすいものではなかったというのも事実である．

WWW は 1989 年に CERN[1] の Tim Berners-Lee 氏らによって提案されたハイパーテキストの概念に基づく分散型情報システムである．主な特徴として次のものが挙げられる．

- Internet に散在する情報をハイパーテキストの概念で整理する．
- ハイパーテキスト情報をやりとりするための手段を提供する．
- さまざまなリソースを指定するための統一的な手段を提供する．

WWW を利用するためのクライアントは数多く開発されており， Internet 上でフリーソフトウェアなどの形をとって配布されている． イリノイ大学の NCSA[2] で開発された Mosaic もその一つである． NCSA Mosaic は 1993 年にリリースされて以来，その完成度の高さや操作性の良さから，現在では WWW の代表的なクライアントとして利用されている．MS-Windows や Macintosh 上で動作する Mosaic も NCSA からリリースされている．

3 WWW のアーキテクチャ

WWW を構成する中心的概念は URL, HTML, HTTP の 3 つである． それぞれ一言でいうならば，URL はインターネット上のさまざまな資源を統一的に表記するための記法，HTML はハイパーテキストを表現するための言語，HTTP は WWW のクライアントとサーバがハイパーテキストをやり取りするためのプロトコルである．

本節ではこれらの 3 つの概念を簡単に説明し，さらにこれらの組み合わせとしての WWW が実際にどのように実現されているのかを示す．

3.1 URL

URL (Uniform Resource Locator) とはインターネット上のさまざまなリソースを統一的に記述するための記法である．一般に，あるリソース

[1] Conseil Européenne pour la Recherche Nucléaire
[2] National Center for Supercomputing Applications

を特定するためにはホスト名, ファイル名, アクセスに用いるプロトコル, アクセス時のコマンドなどさまざまな要素が必要になってくる. URL はこれら全てを取り込んだ名前空間を作ろうというものである.

URL は IETF の URI ワーキンググループで議論が続けられており, 現在は Internet-Draft としてドキュメント[3] が出ている. URL の一般的な書式は次の通りである.

$$scheme://host.domain[:port]/path[?key]$$

URL 中の *scheme* 部には telnet, file, http, gopher, nntp, wais などのさまざまなプロトコルを指定することができる. 具体的な URL の記述例を次に挙げるので参考にされたい.

- ftp サーバ ftp.iij.ad.jp 上のファイル /pub/ls-lR.Z を示す URL :
  - file://ftp.iij.ad.jp/pub/ls-lR.Z
- WWW サーバ www.ncsa.uiuc.edu 上のホームページ index.html を示す URL :
  - http://www.ncsa.uiuc.edu/index.html
- Gopher サーバ gopher.nic.ad.jp (port 70) を示す URL :
  - gopher://gopher.nic.ad.jp/

このように URL を用いることにより, Internet 上のさまざまなサービスから得られる情報を全て統一的に表記し, 取り扱うことが可能になる.

3.2 HTML

WWW は Internet に散在する情報をハイパーテキストとして表現するということは前述したとおりである. このハイパーテキストを記述するための言語が HTML (HyperText Markup Language) である. HTML は SGML[4] の DTD (Document Type Definition) によって定義された言語であり, 次のようなものを表現する能力を持つ.

- 文書のフォーマット情報.
  - タイトル
  - ヘッダ
  - パラグラフ
  - リスト (箇条書き)
  - センテンス / 語句の強調
- さまざまなリソースへのリンク情報.
  - イメージ (静止画, 動画)
  - 音声
  - HTML 文書
  - 他プロトコルのサービス

HTML ではこれらの情報を表現するために SGML 記法に従ったタグを用いる. すなわち<*directive*> ... </*directive*> のような表現形式である. 以下に HTML 文書の簡単な例をあげる.

```
<TITLE>A simple example</TITLE>
<H1>A simple example</H1>
Welcome to <EM>WWW</EM> world. <P>
<UL>
 <LI>
  <A HREF="http://www.fujitsu.co.jp/">
   FUJITSU Ltd. home page </A>
 <LI>
  <A HREF="http://www.ntt.jp/">
   NTT home page </A>
</UL>
```

各タグの詳細は 3.2.1 節以降で説明する. これを WWW のクライアントソフトウェアで表示させると, タグの情報をもとにフォーマットされた文書として見ることができる. 各クライアントの設定により差異はあると思われるが, NCSA Mosaic ではおおよそ次のように表示される.

[3] draft-ietf-uri-url-07.txt

[4] Standard Generalized Markup Language: 文書の論理構造を記述するための言語. 1986 年に ISO8879 として制定.

A simple example

Welcome to *WWW* world.

- FUJITSU Ltd. home page
- NTT home page

上の例で用いたタグはいずれも利用頻度が高く，非常に基本的なものである． タグに用いられるアルファベットは大小文字の区別が無視される． また個々のタグはその綴りから機能が予想できる程度に簡潔である． <TITLE> は文字どおりタイトルを表し，<H1> は見出し，<EM> は文字の強調，<P> はパラグラフの終了，<UL> は番号なしのリスト，<A> はリンクを張るためのアンカーをそれぞれ意味している． 上の例は FUJITSU Ltd. と NTT のホームページへのリンクを含む HTML 文書である． HTML を用いると，この例のようにフォーマット情報を持つハイパーテキストを非常に簡潔に記述することができる．しかし HTML のタグはあくまでも文書の論理構造を表現するものであり，きれいに表示することを目的としたものではないことに注意されたい．

HTML の Internet-Draft は現在期限切れとなっている．しかし IETF の HTML Working Group が作られており，現在もそこで HTML に関する議論が続けられている [8]．

HTML には 3 つのレベルがあり level 1 が最も基本的な部分を定めている．level 2, 3 に関しては一部は使われているが未だ開発途中の段階である [7]．図，表，数式，ユーザ入力などがサポートされている．HTML+ と呼ばれているものは level 3 に該当する． 次節以降では level 1 で定められているタグのうち，いくつかの主要なタグについて説明する．

3.2.1 タイトル

タイトルは基本的に文書の識別子として利用され，WWW クライアント側においては文書本体とは別に表示されることが多い． 一般には HTML 文書の最初に指定される．シンタックスは次の通り．

<TITLE> *text of title* </TITLE>

3.2.2 見出し

HTML の見出しには 6 段階のレベル (1-6) があり，1 が最上位のレベルを意味する．シンタックスは次の通り．

<H*n*> *text of heading* </H*n*>

n は見出しのレベルである．NCSA Mosaic においてはレベルの差異は，見出しのフォントの大きさや太さを変化させることで表現されている．

3.2.3 パラグラフ

HTML 文書においてはタブや改行コードは原則として無視される． また複数個続くスペースも 1 個のスペースとして扱われるようになっている． したがってきれいに整形されたテキストを WWW クライアントに表示させても，期待したような結果は得られない． HTML 文書においてパラグラフの区切りを表現するには <P> というタグを挿入する． WWW クライアントはこのタグに出会ってはじめてパラグラフの区切りを認識する．なお </P> というタグは使われない．

HTML 文書内における空行やインデンテーションは，表示には何の効果も及ぼさないため，基本的には必要のないものかもしれない． しかし HTML 文書のソースの可読性や保守性を保つためには，適当な箇所に空行を入れ，必要ならばインデンテーションを行うことも大切であると思われる．

3.2.4 リスト

ここでいうリストとはいわゆる箇条書きのことである． HTML ではフォーマットの異なるいくつかのリストをサポートしている．それらのうち，3 つの基本的なリストの種類とその書式を次に示した．

- Unnumbered Lists : 通常の箇条書き．LaTeX の itemize 環境に相当する．

```
<UL>
 <LI> List Item A
 <LI> LIst Item B
</UL>
```

- Ordered Lists : 番号付きの箇条書き．LaTeX の enumerate 環境に相当する．

```
<OL>
 <LI> 1st item
 <LI> 2nd item
</OL>
```

- Descriptive Lists : 見出し付きの箇条書き LaTeX の description 環境に相当する．

```
<DL>
 <DT> Description Title
 <DD> Description Description
</DL>
```

これらのリストはネストして用いることが可能である．すなわち，リストの項目がさらにリストを含んだり，あるいは複数のパラグラフを含んでも構わない．

3.2.5 フォーマット済みテキスト

HTML 文書においては改行コード，タブなどは原則として無視されると 3.2.3 節で述べた．例外は <PRE> タグを次のような書式で用いる場合である．

<PRE> *preformatted text* </PRE>

このタグで囲まれた部分 *preformatted text* においては改行コード，タブ，スペースが有効なものとして扱われる．他の HTML タグは通常どおり使用できる．

3.2.6 文字フォーマットと特殊文字

HTML には語句または文をある特定のスタイルで表示させるためのタグが用意されている．次のような書式で用いる．

<*format-style*> *text* </*format-style*>

<*format-style*> タグにはさまざまな種類があるが，ここでは論理レベルのものを紹介するにとどめる．タグの種類と NCSA Mosaic による一般的な表示結果を表 1 に示した．

また HTML においては，ASCII character set のうち “<”, “>”, “&” の 3 つはそれぞれタグの開始，タグの終了，エスケープキャラクタという特殊な意味を持っている．したがって，そのままではそれらの文字自体を表示することはできない．これらの特殊文字の表示には表 2 に示すようなエスケープシーケンスを用いる．

表 1: 論理的なフォーマット指定

タグ	表示	説明
<EM>	*emphasis*	語句の強調
<CITE>	*citation*	引用部分
<VAR>	*filename*	可変部分
<STRONG>	**strong**	強い強調
<code>	code	コード
<samp>	% ls	例示
<kbd>	Home	キーボードのキー

表 2: エスケープシーケンス

ESC Sequence	Symbol
<	<
>	>
&	&

他に，ISO 8859-1 の ü, ñ, à などアクセント記号のついた文字を表示するためのエスケープシーケンスが準備されている．詳細は文献 [6] を参照してほしい．

3.2.7 リンク

HTML では文書内のある部分から他の HTML 文書へ，またはイメージなどへのリンクを構成することができる．リンクはハイパーテキストの根幹をなす部分であり，HTML の数ある機能のうちでも最も重要な機能である．HTML ではリンク先として URL (3.1 節参照) を指定することができる．リンクの簡単な例を次に示す．

```
<A HREF="LinkExample.html">link</A>
```

A は Anchor の先頭文字である．<A> タグに属性としてリンク先を明示することにより，リンクを構成することができる．この例ではキーワード

“link” からローカルディスク上の文書 LinkExample.html へリンクが張られている． これは略記法であり，LinkExample.html の前にクライアントプログラムによって `http://localhost/` などが補完される．

WWW クライアントが文書を表示する際，ハイパーテキストリンクが張られている部分は，リンクが存在することを明示するために何らかの装飾が施される．例えば NCSA Mosaic の初期設定では色を変えた下線付きの文字列として表示される．そしてその部分でユーザが何らかのアクションを起こす (例えばマウスでクリックする) ことにより，リンク先の文書が表示される．

例えば次の例はネットワークを介して CERN のホームページへリンクを張ることを意味する．

```
<A HREF="http://info.cern.ch/">CERN</A>
```

このように URL 記述を用いることにより，リンクの先がローカルに所有する文書であっても，ネットワークを介したリモートホスト上の文書であっても，同様な形式でリンクを構成することができる．すなわちイパーテキストを非常に簡潔に記述することが可能になっている．

3.3 HTTP

WWW ではクライアントとサーバが通信するためのプロトコルを規定している．それが HTTP (HyperText Transfer Protocol) である．HTTP は TCP/IP 上で実現されている．HTML で記述されたハイパーテキストは主にこのプロトコルにしたがってやりとりされる．

HTTP の Internet-Draft も HTML と同様に現在は期限切れとなっている．詳細なスペックは文献 [9] を参照してほしい．

3.4 WWW の動作例

これまで WWW の中心的概念である URL, HTML, HTTP の 3 つを解説してきた． その 3 つの組み合わせである WWW が実際にどのように動作しているのかを簡単に説明する．

一般に「WWW を使う」と言った場合，クライアントに表示された HTML 文書を閲覧し，さらに欲しい情報を求めて文書中のリンクをたどるという動作の繰り返しを指すことが多い．このサイクルを解説するとおおよそ次のようになっている．

1. クライアントは URL 中に書かれたサーバに対して，URL で指定された文書を返送するように要求する．この要求と返送の方法を規定しているのが HTTP である．
2. クライアントは送り返されたデータを適当な方法で表示する． データが HTML 文書であればタグにしたがってフォーマットし，結果を画面に表示する． 画像であれば外部のビューアを起動させることもできる．これはクライアントの設定によって変更できる．
3. ユーザからリンク先の文書が欲しいという要求が出される．例えば NCSA Mosaic ではマウスでクリックするという動作がこれに該当する．
4. 1 に戻る．

このサイクルの繰り返しによって，ユーザは Internet 上の情報を渡り歩くことが可能になる．ここで重要な点は，ユーザは情報がどこにあるのかを，またリンク先の情報はどんな種類なのかを特に意識することなくリンクをたどることが可能であるという点である． これこそが WWW が Internet ナビゲータと呼ばれる所以であろう．

4 WWW によるメディアの統合

WWW プロジェクトによって，ハイパーテキストの概念を用いた Internet ナビゲーションの手段が提供された． Internet 上に散在する情報の間にはリンクが張られ，メディアの形態にとらわれることなく自由にアクセスすることが可能になった．データベースと WWW の連携を行っている例としては USA の Library of Congress[5] がある．すなわち HTML 文書, 画像, 音声, Gopher, WAIS, オンラインデータベースなどさまざまな形態のメディアを, WWW という一つの世界の中で統合して扱うことが可能になったのである．

[5] http://lcweb.loc.gov/homepage/lchp.html

ユーザはWWWを利用することにより，メディアによってアクセス方法を使い分けるという繁雑な作業から解放され，リンク先が何であろうと，「マウスクリック一つ」(NCSA Mosaicの場合)でそこにある情報を享受することができるようになった．

HTML level 2では<FORM>というタグが導入されている．これにより今以上にインタラクティブな要素を取り入れることも可能になっている．例えばPizzaHutのピザ注文用シート[6]は，WWWの応用の一つとして非常に興味深いものがある．対話的なメディアを組み入れることにより，このようにアイディア次第で色々な応用が考えられるようになっている．

今のところWWWは，さまざまなメディアを取りまとめるための共通基盤として大きな可能性を持っていると言えよう．

5 WWWの問題点

分散型情報サービスとしてWWWは急激な広がりを見せつつある．そのクライアントプログラムとして有名なNCSA Mosaicは以下のような特徴を持っている．

- すぐれたUIを持っており非常に使いやすい．⇒ 現在ではWWWクライアントの事実上の標準となっている．
- マルチメディア化されているため画像や音声などを容易に扱うことができる．

しかし，この特徴ゆえにネットワークのトラフィックを容易に増大させてしまうという問題をひき起こしている．NSFNETを流れる全パケット数は1993年8月から1994年8月の間におよそ2倍に増えているが，WWWのパケット数だけを見ると約30倍にまで増加している．比率で見ると全パケット数のうちWWWパケットの占める比率は0.37%から5.2%へと増大している[3]．

WWWによるトラフィックの増加を軽減させるための一つの手段として，以前にアクセスされたデータをネットワーク的に近いWWWサーバでキャッシュしておくという対策がなされている．

他に現時点で検討を要する点としては次のようなものが考えられよう．

- メディアの同期再生機構がない．
- 国際化に関すること．
- ハイパーテキストの編集環境が必要．

これらは，いづれもクライアントのインプリメントや規格自体を改良することで解決できるものと考えられる．興味のある方はWWWに関するWGやメイリングリスト，国際会議などの場に積極的に参加して欲しい．

6 おわりに

WWWはInternetという情報のるつぼを自在にナビゲーションするための手段を提供し，さらに画像や音声など複数のメディアを扱えるようにしたという点において画期的である．さらにCGIインターフェースの導入により，従来から使われているキャラクタベースインターフェースを持つデータベースシステムのフロントエンドとしても十分実用的に使えるものとなっている．それゆえ全世界で広く受け入れられているのであろう．

WWWのような分散型情報システムの登場により，最近は初心者がInternetに対して感じるしきいが低くなっている．Internetが全世界に急速に広まりつつあるが，WWWはその要因のうちの1つに数えられると思われる．

WWWはまさに今，その開発が行なわれている最中である．未だ問題点も存在しているし，規格自体が審議中の部分もある．　またWWWにはさまざまな利用価値が考えられ，今後どのような形態で発展 / 利用されていくのか予測不可能な部分が大いにある．　さまざまな可能性を秘めたWWWは今後のInternetの方向性，ひいてはネットワーク社会の方向性に大きな影響を及ぼすものであると筆者らは予感している．

参考文献

[1] NSFNET History of hosts, available from

[6] http://www.pizzahut.com/

ftp://nic.merit.edu/nsfnet/statistics /history.hosts.

[2] NSFNET History of Usage by Services, available from ftp://nic.merit.edu/nsfnet /statistics/history.ports.

[3] NSFNET Backbone Statistics Page, available from http://www.cc.gatech.edu/gvu /stats/NSF/merit.html

[4] A Beginner's Guide to HTML, available from http://www.ncsa.uiuc.edu/General /Internet/WWW/HTMLPrimer.html

[5] A Beginer's Guide to URLs, available from http://www.ncsa.uiuc.edu/SDG/Software /Mosaic/Demo/url-primer.html

[6] ISO latin 1 character entities, available from http://info.cern.ch/hypertext/WWW /MarkUp/ISOlat1.html

[7] HyperText Markup Language (HTML), available from http://info.cern.ch/hypter- text/WWW/MarkUp/MarkUp.html

[8] HyperText Markup Language (HTML) Charter, available from http://info.cern.ch /hypertext/WWW/MarkUp/HTML-WG /Charter.html

[9] HyperText Transfer Protocol (HTTP), available from http://info.cern.ch/hyper- text/WWW/Protocols/HTTP/HTTP2.html

[10] 高田敏弘： World Wide Web の日本語化・国際化の現状, IAJ NEWS, Vol.1, No.2, pp.15-21 (1994).

[11] 齋藤正史, 山口英 ： インターネットの情報サービス, 情報処理, Vol.34, No.12, pp.1415-1421 (1993).

[12] 小町祐史 : マルチメディア / ハイパーメディア情報交換の標準化動向, 情報処理, Vol.35, No.7, pp.632-641 (1994).

[13] 吉村伸 : インターネットの利用と仕組み, UNIX MAGAZINE, 1993/12, 1994/2, 3, アスキー.

[14] 新たな社会基盤となるマルチメディア網の本質, 日経コンピュータ, 1994.6.13, 日経 BP.

[15] 急速に拡大する Internet ビジネスツールとして急浮上, 日経バイト, 1994.7, 日経 BP.