

診断型 WWW 情報検索システム (2) 構成と実装*

1 A C - 5 富沢伸行 二瓶克己 柴田晃宏 島津秀雄†
NEC C&C メディア研究所†

1 はじめに

問題解決事例やノウハウ情報等を、多くの人々が容易に検索・活用するためのシステムとして、我々は、診断型 WWW 情報検索システムの開発を行なった。本稿では、その構成と実装について報告する。

2 診断型 WWW 情報検索システム

我々の診断型 WWW 情報検索システムは、我々が開発したヘルプデスク業務支援システム Help Desk Builder[1],[2],[3]の知識検索部 BT (以下 BT と略記)をベースに WWW 上に実装、さらに自動質問による絞り込み [5]の機能を実装したものである。

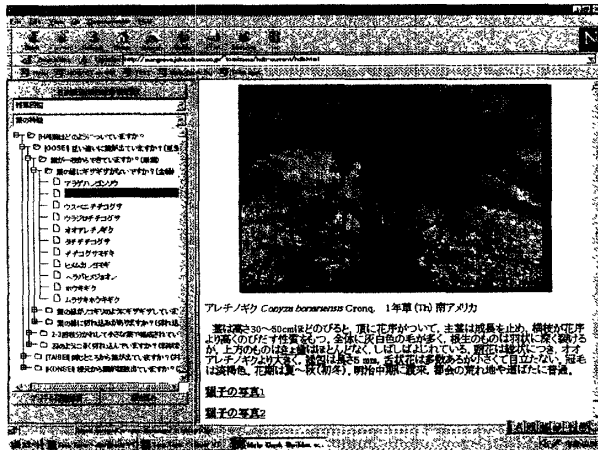


図 1: 診断型 WWW 情報検索システム画面例: 左が目次表示部、右がページ表示部

BT は「ハイパーブック」を単位として情報を整理/管理する。一つのハイパーブックの中には、「目次」と呼ばれる複数の階層構造インデックスと、相互にリンク付けされた複数の「ページ」が存在する。「ページ」は相互に関連付けられたハイパーテキストシステムであり (図 1 参照)、さらに、各ページは目次中の各々の項目からも関連付けられ、複数の目次を使用してページを分類することができる。BT では、ハイパーブック中のページに対して、キーワード検索、フリーワード検索に加え

*A Diagnostic-based Information Retrieval System on the WWW(2), Architecture and Implementation.

†Nobuyuki TOMIZAWA, Katsumi NIHEI, Akihiro SHIBATA, and Hideo SHIMAZU

†NEC C&C Media Research Laboratories., NEC Corp.

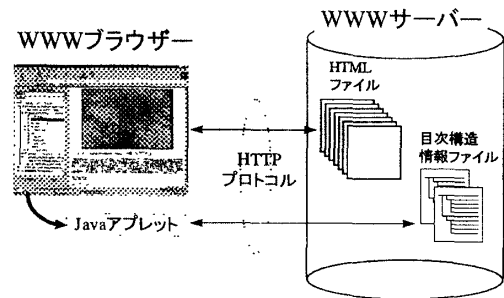


図 2: 診断型 WWW 情報検索システムの構成

て、目次の構造を検索のキーとした目次検索や、さらにこれらの各種検索機能を and/or で組み合わせることができる統合検索を持つ。このシステムの機能の詳細については、[1], [2]等を参照されたい。

3 設計方針

本システムは、[4]で述べた方針に基づき、WWW ブラウザー上に Java アプレットを用いて実装することとした。そのため、本システムを利用する上で WWW ブラウザー以外のアプリケーションをインストールする必要がない。また、文書フォーマットとして HTML 形式を使用するためデータの交換も容易になる。さらに、WWW システム上の CGI 機能を利用することで、BT のデータベース検索機能も容易に実現可能である。

4 構成

以上の方針をもとに開発した診断型 WWW 情報検索システムは図 2 のような構成となっている。

BT は、複数目次と各種の検索機能を持つハイパーテキストシステムとみることができるが、このハイパーテキスト部には WWW の枠組をそのまま利用する。つまり、BT のページが WWW のページに対応し、BT におけるリンクも WWW の URL によるリンクをそのまま利用する。これによって、BT が管理するページと通常の WWW ページを同等に扱うことができる。

BT 独自の構成要素である目次やページ属性 (キーワード等) の情報を利用する機能を WWW システムに

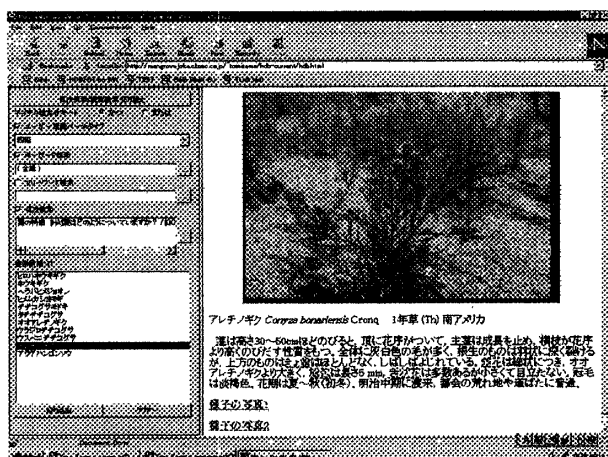


図 3: 診断型 WWW 情報検索システム画面例: 左が統合検索パネル、各種検索を組み合わせて検索可能

付加するために、クライアント側ではアプレットを、サーバー側では CGI を利用して実装することにした。このような構成にすることで、既存の WWW サイトに対し BT の検索機能を容易にインストールすることができる。

5 実装

以下では、現在の実装に関して簡単に述べる。

サーバー側 ページに相当する HTML ファイルは通常の WWW と同様に処理すればよいので、サーバー側では特別な処理をする必要はない。また、目次の情報も通常のファイルとして、HTTP プロトコルを介してクライアント側に転送されるため、これも特別な処理は不要である。ただし、全文検索の処理は、クライアント側からの要求に応じて、CGI を通してサーバー側で行なう。

クライアント側 ページの表示は WWW ブラウザーを利用するために特別な処理は必要無い。Java アプレットによって実装されているのは、フリーワード検索機能を除く全ての検索機能と目次表示部(図 3 参照)である。

目次表示部の操作によりページ表示操作が必要になった場合、アプレットから指定した URL の WWW ページを表示させる Java の API を発行するだけなので、その処理はごく簡単なものである。

フリーワード検索を行なう場合、各種検索エンジンによる検索と同様に、検索のための条件式をクライアント側で作成、HTTP プロトコルでサーバー側の CGI プログラムを呼び出して、その検索結果をアプレットで受け取り、処理する。

フリーワード検索以外の検索(例えばキーワード検索)を行なう場合、基本的にはすべてクライアント側で検索を行なっている。これは、Java アプレット上で検索処理を行なってもある程度の検索速度が得られると考えられること、逆に一回の検索を実行するたびにネットワーク越しに要求を出してサーバー側で検索を行なうと、検索速度が WWW サーバーの負荷状態や回線速度に依存してしまうためである。このように、クライアント側で検索処理を行なうことで、基本的にはサーバーとの通信は生じない(目次の構造等の情報は全て必要になった時にロードされるので、そのような情報が検索時に必要になった際に、まだロードされていなければ、その情報の読み出しは行なわれる)。

現在、実装で問題なのは起動時の処理(アプレットの読み込み)と、特定の検索機能の初回起動時に時間がかかる点である。今後これらの点の改善をすすめていく。

6 おわりに

本稿では、現在開発中の診断型 WWW 情報検索システムの構成と実装について述べた。WWW 上の枠組を利用することで、多くの環境で利用できるシステムを比較的効率良く実装することができた。現在、BT のアクセス制御機能の WWW 上で実現する枠組を検討中である。今後、使用ユーザーやアクセス権に基づいて目次やページの表示を可変にする機能を実装する予定である。

なお、本文中の画面例は、農林水産省 草地試験場 黒川俊二氏が作成したものを、筆者らが本システム向けにデータ変換したものである。データの作成者である黒川氏に感謝する。

参考文献

- [1] 二瓶ほか, “ハイパーテキストドキュメントによる事例ベース検索システム”, 人工知能学会研究会資料 (SIG-KBS-9502-7), pp.48-pp.55, 1995.
- [2] 柴田ほか, “ヘルプデスク向け情報ベース構築ツール”, 人工知能学会研究会資料 (SIG-J-9501-19), pp.128-pp.135, 1995.
- [3] 長妻ほか, “ヘルプデスク構築支援システム「Help Desk Builder™」の開発 問い合わせ管理 -Help Desk Builder/CS-”, 情処 54 全大, 1997.
- [4] 柴田ほか, “診断型 WWW 情報検索システム (1), — 開発方針 —”, 情処 55 全大, 1997.
- [5] 二瓶ほか, “診断型 WWW 情報検索システム (3), — 自動質問 —”, 情処 55 全大, 1997.