

1 エージェントによる生産支援技術の開発

村河亮利, 藤原義也, 岩谷 稔, 田尾博幸, 吉野信行

Development for manufacturing support technology with agent

MURAKAWA Akitoshi, FUJIWARA Yoshinari, IWATANI Minoru, TAO Hiroyuki and YOSHINO Nobuyuki

In present manufacturing industry it is required high technology, quick production capacity and information collection. However it takes too much time to reach the information of purposes, when engineers gather information using the Internet. Since the home pages are created, are renewed and are deleted, information that has been collected in having point of time becomes old with time. So we make a prototype of a system that extracts results from automatically gathered information from the Internet and created databases.

In addition, the extraction of the table considering the unit system becomes possible.

キーワード：エージェント，ネットワーク，インテリジェント，情報収集

1 緒 言

製品開発時に検索エンジンを利用した情報収集は主に手動で行われている。そのため利用者は時間と労力を要しているのが現状である。そこで、利用者が必要な情報の条件設定を行えばインターネット上から自動的に情報収集・提示することにより、利用者の負担を軽減するシステム開発を行ってきた。本年度は、電子回路設計の効率的なデータ収集に対し単位まで考慮したデータ抽出や、余分なページを参照しないアルゴリズム等を提案し適用した。また、製品・サービスへの適用を図ることで新市場の開拓支援を行う。

2 シ ス テ ム

システム概要およびシステムの構成について述べる。

2.1 システム概要

現在、分散データベース（以下 DB）から目的とするデータを収集するシステム¹⁾²⁾や検索エンジン³⁾は開発されている。しかし、時々刻々と更新されるインターネット上のデータを収集し、利用者が望む情報を整理し提示するシステムは見当らない。そこで、昨年度作成したデータ収集システム⁴⁾を使用し、インターネット上で広く使用されている Hyper Text Markup Language(以下 HTML)中から目的のデータを抽出・加工し、利用者に提示する手法について検討する。具体的には、まず利用者が検索エンジンを用いてキーワードを入力し検索結果アドレスを対象に検索を行うか、検索対象のアドレス(以下 URL)の指定を行う。また必

要に応じて独自に作成した DB から情報を引き出す。引き出したアドレス情報から利用者が必要とするデータと一致した情報のみを提示することにより、利用者の作業負担を軽減し、効率よくデータ収集が可能なシステムを構築する。本年度は電子回路設計に必要なデータ収集を行い、システムの検証および問題点の抽出を行う。

2.2 システム構成

全体のシステム構成を図1に示す。利用者は Web ブラウザを通してキーワード 要求項目等を入力する。入力データは Web サーバを通じて検索サーバに送信され、インターネットや独自データベースから情報収集を行う。収集した情報から、利用者の要求項目とデータが一致するものを選択し、Web サーバを通じて利用者に提示する。

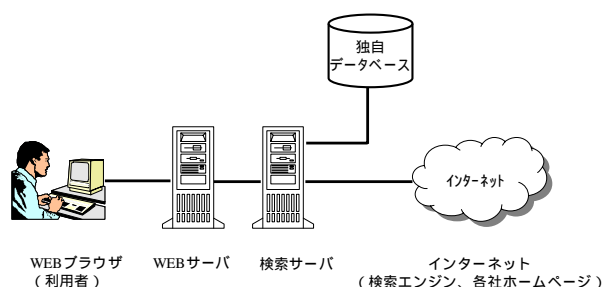


図1 システム構成

3 検索方法

昨年度の検索システムに機能追加（単位考慮，ハードウェア記述言語（以下 HDL）の抽出，検索エンジンとの連携，リンク登録）および入力部の改善を行った。

3.1 単位考慮

昨年度までは，表データを抽出し，項目とその内容が文字列で一致する場合のみデータ抽出・提示する機能しか備えてなかった。本年度はこの機能に加え，収集した表データから，単位まで考慮しデータ抽出を行った。例えば，項目を CPU として 100MHz 以上 200MHz 以下とした場合表データに 150MHz とあれば抽出し，50MHz とあれば抽出を行わない。これを実現するための手順を次に示す。

単位として認識できるのは 1 次元 (m, Hz 等) と，単位に付く接頭語 (G, k, μ 等) である。これらを構造体に格納し，単位が一致した数値の大小比較を行う。単位であることを判断する状態遷移図を **図 2** に，格納する数値データの構造体を **図 3** に示す。

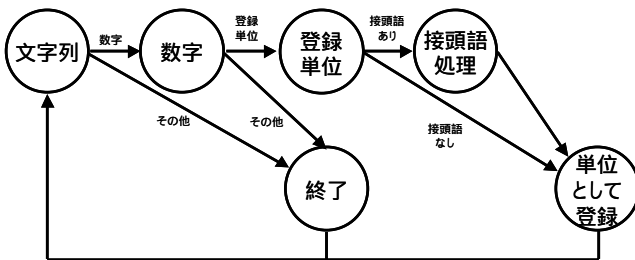


図 2 状態遷移図

表 1 数値データの構造体

要素名	例
数値	100
単位	Hz

3.2 ハードウェア記述言語(HDL)プログラム抽出

製品開発の電子回路設計時に必要な HDL ライブラリを，公開されているホームページから抽出した。

抽出対象として HDL プログラム言語の 1 つである Verilog 記述を選択した。この言語の記述方法は特定文字列 (module) で始まり特定文字列 (endmodule) で終了する。ここで収集したホームページの中から HDL ライブラリを，検索対象のホームページから上記の特定文字で囲まれる部分の抽出検証を行った。 **図**

3 に HDL プログラム抽出例を示す。

```

module aoi(a,b,c,d,f);
    input a,b,c,d;
(中略)
    assign f = ~0;
endmodule
  
```

図 3 HDL(Verilog)プログラム抽出例

3.3 検索エンジンとの連携

検索エンジンを複数選択可能とした。また，検索結果のソースの規則性を利用し，新たな検索エンジンを簡単に追加できるようにした。例えば「検索結果のソースの<!--RESULT-->と<!--RESULT-->タグで囲まれた部分に検索結果の URL がある」がある。このような共通点を整理し **表 2** に示すような 17 項目に分類することにより，新規に検索エンジンを加える場合や，検索エンジンの出力フォーマット等が変更になった場合，**表 2** の内容をテキストファイルにすることで，実行プログラムを変更させることなく対応可能とした。

表 2 検索エンジン対応パラメータ (抜粋)

1	検索エンジン名
2	検索結果を表示しているページの値 (例: 0)
3	1 ページに表示する検索結果数 (例: 20)
(中略)	
15	表示ページの初期化
16	検索数の存在部分の始点
17	検索数の存在部分の終点

3.4 リンク登録

昨年度のシステムでは検索エンジンでヒットしたサイトのディレクトリ内にあるリンク先を無制限に検索していたため，関係のないページまで検索し時間がかかるという問題と，必要なデータは存在しないと考えられる深い階層にある不必要なリンク先まで調べてしまうという問題があった。

そこで今年度は利用者が検索するリンク階層の深さを指定可能とした。URL 情報を保存する DB のフォーマットは **図 4** のようになっており，左からカンマ区切りでリンク先情報，URL の情報源，リンク階層の情報が含まれている。リンク先情報には検索対象となる URL が格納され，URL の情報源にはリンク先情報を取得した情報源が格納される。例えば **図 4** の 1 行目では「http://xxx.yyy.zz/fpga/」は検索エンジン「goo」から取得した URL であることを示し，6 行目の「http://xxx.yyy.zz/fpga/freeip/index.html」は「http://xxx.yyy.zz/fpga/」のページに含まれていた URL であ

ることを示す。URL の情報源は、例えば3.2に示すようなHDLライブラリを検索した場合、ページにプログラムのみが書かれており、そのプログラムの説明がそのプログラムのURLのリンク元に書かれている場合があるため、リンク元のURLを参照可能とするためにDBに保持している。

また、リンク階層は検索エンジンでヒットしたURLを0階層目として、0階層目に含まれるリンク先を1階層目、1階層目に含まれるリンク先を2階層目（以下同様）としている。

URLの検索はURL情報を格納するDBに納められた順にホームページの内容を調べていき、利用者が指定したリンク階層に到達するまでURL情報のDBに追加され、DB内のリンク先をすべて調べ終わると検索終了となる。

"http://xxx.yyy.com/fpga", goo, 0
 "http://www.xyz.co.jp/a/b26/26-7.8.html", infoseek, 0
 "http://www.yyy.edu/fpga/fpga.html", fresheye, 0
 (中略)
 "http://www.zzz.edu/ee/fpga2003/", fresheye, 0
 "http://xxx.yyy.zz/fpga/freeip/index.html", http://xxx.yyy.zz/fpga/, 1
 "http://xxx.yyy.zz/fpga/soho/index.html", http://xxx.yyy.zz/fpga/, 1
 (以下略)

図4 URL DBのフォーマット

3.5 入力部

本システムはWebブラウザからの操作が可能である。操作入力画面を図5に示す

入力ページは一般に利用されている検索エンジンの入力部に近いものとした。これにより利用者が違和感なく利用できることを目標とし設計を行った。このページには詳細な設定ができるよう入力部を多く配置しており、デフォルトの状態では検索キーワードと抽出項目を入力するだけで情報の収集が可能である。図5の詳細を以下に記す。

1. 検索エンジンを利用して「キーワード」を入力しURLを抽出する。キーワードのAND検索のみ可能である。
2. 「URL数」は検索エンジンにより抽出するURLの数を指定する。検索エンジンからの上位の結果のみを検索する場合、この数を少なくすることにより調節可能である。
3. 「抽出項目」はURLのソースより指定項目を抽出する。数値を入力する場合は「単位」、「抽出範囲」を指定することが可能である。
4. 「検索エンジン」の項目にチェックをすることで

URLを収集するとき利用する検索エンジンを指定する。



図5 操作入力画面

4 結果および考察

実行環境および検索条件を設定し実行した結果を以下に示す。

4.1 実行環境

実行環境（検索サーバ）を表3に、実行条件を表4に示す。

表3 実行環境（検索サーバ）

CPU	Pentium 2.8GHz
RAM	2GB
OS	Windows2000
開発言語	Java
回線速度	100Mbps(ベストエフォート)

表4 実行条件

使用した検索エンジン	goo, yahoo infoseek, fresheye
検索件数	各検索エンジン結果の上位100件
リンク階層	5階層
再試行回数	5回

4.2 検索条件および結果

抽出条件を表5に、実験結果を表6、図6に示す。

表5 抽出条件

キーワード	FPGA
-------	------

項目	内容
FLEX	100MHz 以上
MAX	90MHz より大きい

表6 検索ページ数および検索時間

検索したホームページ	9190 件
検索時間	5 時間 44 分

```
http://www.AAA.co.jp/ucnt1.html
FLEX MAX
178MHz 140.84MHz
-----END OF http://AAA.co.jp/ucnt1.html
```

図6 抽出結果(一部)

これらの結果より、1次元(図6ではHz)ではあるが文字列が一致するだけでなく、単位まで考慮して利用者に提示することが確認できる。

4.3 文字列抽出

インターネット上から HDL ライブラリを抽出した条件を表7に、その結果を図7に示す。

表7 抽出設定条件

キーワード	FPGA
開始文字	module
終了文字	endmodule

```
module MEMORY_CONTROL(CLK, RST, .....);
input CLK, RST, RDSTB;
(中略)
endmodule
-----END OF http://www.BBB.co.jp/DDD/Verilog/CCC.v
```

図7 プログラム抽出結果

図7より, Verilog 記述のプログラムが抽出されていることが確認できる。

5 結 言

昨年度のシステムに機能追加(単位考慮, HDL の抽出, 検索エンジンとの連携, リンク登録)することにより, より利用者の要求するデータに適合するものを選択的に抽出可能となった。具体的には, 以下の4点が挙げられる。

- (1) 数字や単位等を考慮し数量の大小を比較可能とした。
- (2) HDL プログラムの抽出を可能とした。
- (3) 各検索エンジンに対応し設定ファイルを用意することにより, プログラム本体の変更をなくし汎用性の高いシステムとした。
- (4) 不必要なリンク先を検索しない機能の追加を行った。

今後, 利用者の意見を取り入れ改善を行ったり, 複雑な単位系(例: m/s^2)に対応する予定である。

文 献

- 1) 西野, 赤坂, 小泉: 情報処理学会論文誌 40(1999)1, 253
- 2) 西野, 木室, 赤坂: 情報処理学会研究報告 98(1998)8, 201
- 3) 森川, 深澤, 吉岡, 大須賀, 本位田: 情報処理学会講演論文集 3(2000), 367
- 4) 村河, 打田, 中塩: 広島県立西部工技研究報告, No44(2001)