

インターネット経由で DIALOG を検索する際の利用者インターフェース

6 L - 5

— 適合度順出力機能を中心として —

相良 佳弘*) 安形 輝*) 高橋 兼清 細野 公男**)

*)慶應義塾大学大学院文学研究科、**)慶應義塾大学文学部

1. はじめに

DIALOG は、従来からの電話回線に加え、インターネットを介して telnet 接続による情報検索サービスを公開し始め、より多くの利用者に門戸を開いた。しかし、その利用者インターフェースはデータベースのコマンド入力形式のままであり、利用者の多くにとっては障害が大きいと考えられる。そこで本稿では DIALOG を利用する際の利用者インターフェースの改善を行うために、図1のような利用者と DIALOG の間に位置するフロントエンドシステムを構築する。

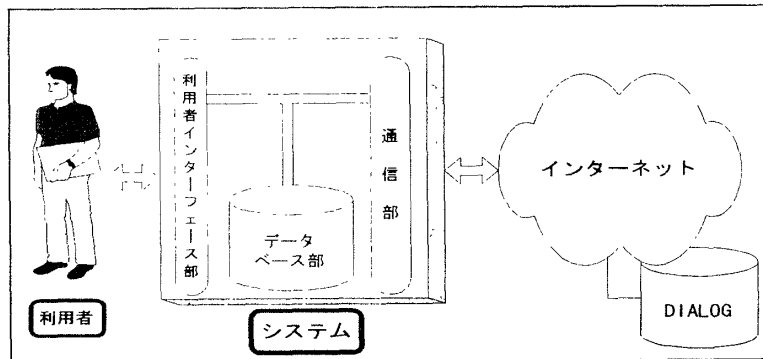


図1 DIALOG フロントエンドシステム

本システムは、サーチャーの代行が一般的である DIALOG の検索を「利用者」にも利用可能にすることを目的としている。ここで想定する「利用者」とは WWW(World Wide Web) 上の Infoseek に代表されるサーチエンジンを検索できる程度の知識と経験をしかない利用者のことである。

2. インターフェースの問題と対応策

利用者が DIALOG を利用する際に大きな問題となるのが、(a)複雑なコマンドの入力、(b)論理演算子による検索式の構築、(c)検索集合の操作と考えられる。そこで、本システムはそれぞれの問題に対応して、以下のことを実現する。

- (a)全体として GUI(グラフィカルユーザーインターフェース)を採用することと、検索から出力の一連の流れを自動化することにより、コマンド入力を減少させる
- (b)検索語入力からの「検索式自動生成」を行う。

つまり、利用者は複雑な論理演算子を扱う必要がない

- (c)検索結果の「適合度順出力」を実現する。

3. 適合度順出力

原則的に DIALOG の検索結果の出力は、出版年の逆順で、さらに、レコード番号順である。これを適合度順出力

に変換するアルゴリズムとしては、基本的には Quorum 検索(レベル別検索)と呼ばれる手法を採用する[1]。さらに、補完的に検索語の出現頻度情報を使った並べ替えを行う。

Quorum 検索は、検索語群を使い、非常に制限の強い検索式から最も緩い検索式までを順に検索し、制限の強い検索式の検索結果から順に出力することで適合度順出力を実現する手法である。例えば検索語 A、B、C (順序は入力順とする)があった場合には、

- S 1 :S A and B and C [Sは集合番号]
- S 2 :S A and B not S 1
- S 3 :S A and C not (S 1 or S 2)

Front-end User Interface System on DIALOG via Internet
 Akihiro SAGARA*, Teru AGATA*, Kanekiyo TAKAHASHI,
 and Kimio HOSONO**
 <{sagara, itasan, keen, hosono}@slis.keio.ac.jp}>
 *Graduate School of Library and Information Science, Keio University
 ** Faculty of Letters, Keio University
 2-15-45 Mita, Minato-ku, Tokyo, 108, Japan

という形で順に検索式を構築し検索を行う。出力はS1、S2、S3の順で行う。

Quorum 検索は明らかに非常に荒い適合度順出力であり、同順位の文献が数多く出力してしまう。そこで、そのような文献の順位付けを行うために、補完的な手法として語の出現頻度情報を採用する。具体的には、同順位の文献を、検索語の出現頻度が多い順に並べ替え出力する。

Quorum 検索を行わず、単に出現頻度情報のみで十分ではないかと考えられるかもしれない。しかし、出現頻度情報を処理するには、抄録データが必要であり、Quorum 検索を行わない場合、時間的に多くのコストが費やされることになる。

4. 試作システムの概要

4.1. 設計・構築

試作システム構築のためのプログラミング言語はその汎用性で注目を集める JAVA とした。設計・プログラミング共にオブジェクト指向技術を取り入れ、構成要素として利用者インターフェース部、データベース部、通信部を分離させた。分離により、各構成要素の他への依存が最小限になり、構

成要素の交換が可能となる。

扱うデータベースは図書館・情報学分野のデータベース LISA (File:61) とする。

4.2 基本動作

試作システムは、起動時に DIALOG に接続しデータベースとして LISA を選択する。利用者によって検索語群が入力されたら、単語に切り分ける。検索語それぞれについて、辞書を使用し語形変化を統一し、さらに不要語を除去する。次に前述の適合度順出力の方法に従い検索を行い、検索結果を出力する。

著者名、書名だけの簡略表示をマウスでクリックすることでより詳細な書誌データを表示するという機能も付与している。

5. 試作システムの動作例

試作システムの動作例を図2に挙げておく。この場合には検索語として"QUORUM"と"SEARCH"という語を使い、検索を行い、その後一番目について詳細を表示したものである。

6. まとめ

ここでは DIALOG を利用するためのフロントエンドシステムを構築した。今後はこのシステムの有効性を検証していくとともに、さらに以下のような機能を拡充していく予定である。

- 1) 適合情報を利用した適合フィードバック
- 2) SCI、SSCI と組み合わせによる自動引用検索あるいは引用分析
- 3) 他データベース、他システムを含めたシステムレスな横断検索

【引用文献】

- [1] Cleverdon, C. Optimizing convenient online access to bibliographic databases. Information services & use, vol. 4, p. 37-47(1984)

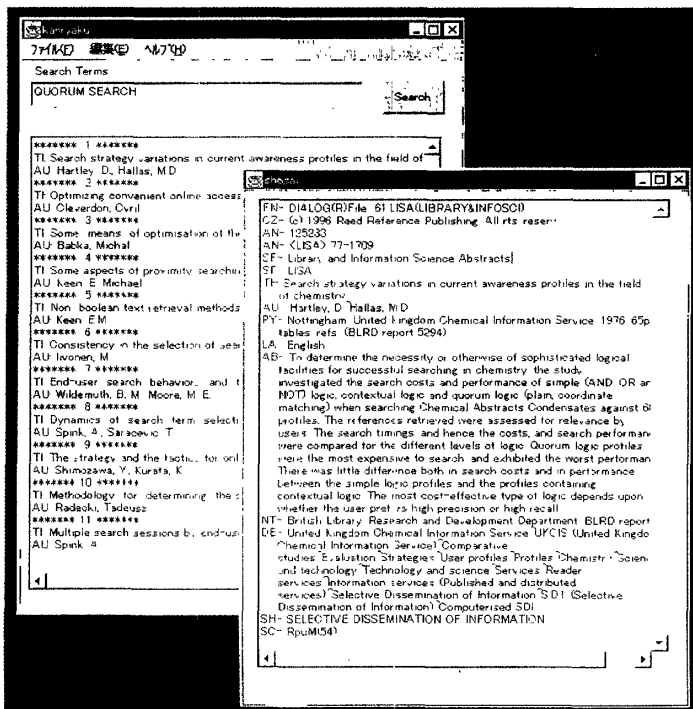


図2 画面例