

# 分散型情報検索システム TROIKA と その応用 (1)

3H-1

平木 しげ子, 荒木 雄豪 (京都産業大学), 今井 恒雄 (富士通株)

## 1. はじめに

現在学術情報センタを中心とする全国ネットワークの整備が進められており、研究者にとって各種データベースの利用が便利になりつつあることは誠に喜ばしいことであるが、現在の段階では利用者が同時に利用できるデータベースは一つずつに限られている。しかしながら研究によっては数個のデータベースに同時にアクセスして検索結果を得たい場合も生ずることがある。このような問題を解決する一つの試みとして異機種、異種データベースに同時接続し、統一されたコマンドで検索できる分散型情報検索システムを試作、実験し、前回の情報処理学会において「異機種上の複数異種検索システムを一つのシステムとして扱うための試み」のテーマで中間発表したがそれ以後の経過を報告する。

一般に異なるメーカーの異なるデータベースを同時に接続し、利用するには次のような項目について十分に検討しておくかねばならない。

第一は異なるメーカーのコンピュータを接続するための通信回線と、そのプロトコル、第二はコード体系(漢字コード系も含む)、第三は応用プログラムが回線を扱う時のアクセスメソッド、第四は情報検索システムのコマンド体系である。

このような項目について、各コンピュータに対して共通の仕様と構造を基本的に満足させ、各コンピュータに組み込むプログラムを出来るだけ共通にし、しかも簡単な論理

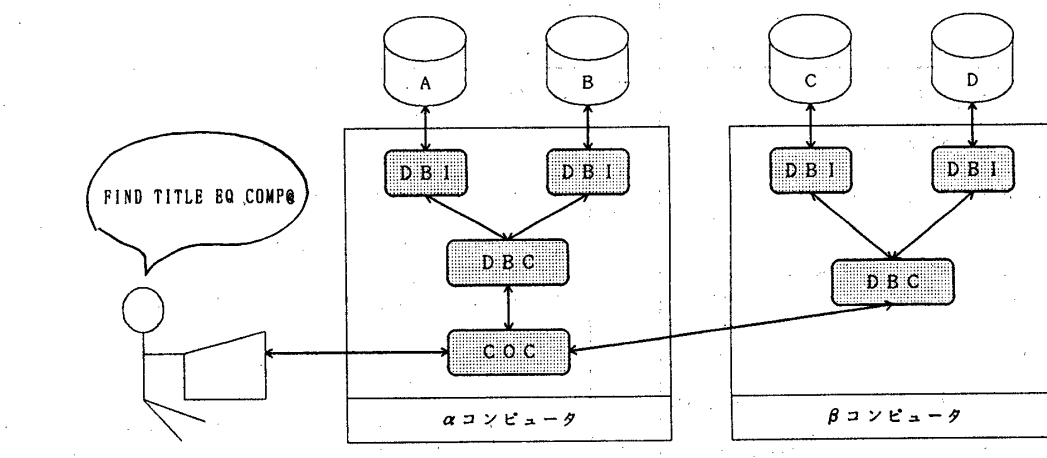
でかつ容易にプログラムを開発できることが、最低限の必要条件であろう。

## 2. TROIKAシステムの特長

このような必要条件を満足させるシステムを目指して TROIKAを開発したわけだが、その特長について述べる。まず通信回線は、プロトコルについてはTTY手順(X-28、X-29も含む)、NI系ネットワークのNVTプロトコルの二つが、大学間という環境下で共通に使える手順であり、これをベースとした。第二のコード系については、英数字カナ、漢字をサポートしているが、TTY手順でのサポートが基本となるための、漢字コード、漢字シフトコードについても配慮をしている。第三のアクセスメソッドについては、TSS環境下が必須の条件となるので、TSSのアプリケーションプログラムが動作して他のコンピュータとNIプロトコル(NVT)によって接続する方式をとっている。第四のコマンド体系に対しては統一コマンドを用いている。

## 3. TROIKA (分散処理) の構成と機能

TROIKAは、図-1にあるようにCOC、DBC、DBIの3つの部分から構成され、COCは複数のコンピュータを使う場合、DBCは一つのコンピュータで二つ以上の検索システムを使う場合、DBIは検索システムごとに組み込まれる。



A distributed information retrieval system TROIKA and its applied system 図-1 システム 概念図

Shigeko Hiraki<sup>1</sup>, Hugo Araki<sup>1</sup> and Tsuneo Imai<sup>2</sup>

1; KYOTO SANGYO Univ. 2; FUJITSU Ltd.,

これら三つのモジュールは基本的にどのコンピュータでも論理的には同じ機能であるのでプログラムもできるだけ共通化して開発した。

各モジュールに共通している重要な機能は、複数のコンピュータ間の通信をできるだけ共通にまた簡潔にしたこと、つまり実端末が接続されているシステムが、利用者インタフェイスを受け持ち、他のモジュールは複数のコンピュータにまたがっても、送受信するデータはすべて共通になるようにしてあることである。

またデータの流れの制御については他のコンピュータのエミュレータとして動作すること、エミュレータ動作で対象コンピュータの切り換えを行うことの二つを実現するために“////I0=”という文字列を使っている。これはあるコンピュータでの処理が終わって次の処理を待つ状態（入力待）を表す。

4. TROIKAでの新規追加機能

これまでのTROIKAに次のような機能を新たに追加した。

(1) DBIの自動作成

DBIはデータベースシステムの違いとデータベースの内容によって、すべて異なってくる。新規にデータベースを追加する時にDBIを次のような手順で自動作成する機能を追加した。データベースの種類ごとに用意した標準プログラムと、対象となるデータベースのDDLの定義体を入力として、そのデータベース用のDBIを自動的に作成する。これによってデータベースの追加やDDLの変更に對して柔軟に對応できるシステムとなる。

(2) DBIで検索システム固有コマンドのサポート

これまでは、固有コマンドは“% コマンド名”と入力したが、操作性を考慮し、予めDBIごとに固有コマンドを定義しておき、操作上では統一コマンドと同じように使えるようにした。

(3) 出力処理での機能強化

a. 表示での見やすさへの配慮

コードを文字に変換して出力する。例えば“1”ならば男、“2”ならば女と表示してデータの見やすさを実現した。この切り換えは実行中に任意に切り換えることができる。

b. ユーザプログラムでの任意フォーマット出力

これまでの統一された7種類の出力様式に加えてユーザプログラムでの自由な出力を可能にした。そのために一件の出力に対して一論理レコードという単位にまとめてユーザプログラムに渡す方式をとっている。

5. おわりに

複数コンピュータにまたがる応用システムとして、TROIKAを例に、必要な要件を実現する方法について述べたが、現状では多くのOSではその構造上、このようなシステムは非常に作りづらいのが現実である。こうした開発環境はなかなか改善されないが、同一メーカーの複数コンピュータでの応用システムを開発する場合には、基本的な構造、モジュールは生かしながら同一メーカーである利点を生かし、汎用性、共通性の面では劣るが操作性のよさ、機能の豊富さで優るシステムを開発できる。その一つの応用例として富士通(株)関西システム部で開発中のシステムを「分散型情報検索システムTROIKAとその応用例(2)」として今回発表する。

表-1 TROIKAテーブル一覧

No.	テーブル名	内 容	所有区分
1	TBLNAME	テーブル名一覧	サブリン
2	CONTROL	ホスト名, 回線番号 etc.	COC
3	BASE	使用可能 DB 名 etc.	DBC COC
4	SET	実行時に参照される変更可能パラメータ	DBI DBC
5	DATABASE	データベース項目名, 属性, etc.	DBI
6	COMMAND	統一コマンド一覧	サブリン
7	BNZAN	使用可能演算記号, 内部演算記号, 対応表	サブリン
8	CURRENT	最新の有効検索コマンドと件数	DBI
9	RESERVE	リザーブ名一覧	サブリン

No.	テーブル名	内 容	所有区分
10	MATCH	現集合のID	DBI
11	ID	コマンドで指定されたID	DBI
12	IMAGE	イメージ表示のための情報	DBI
13	CCHANGE	統一コマンド→個別DBの コマンド	DBI
14	DCHANGE	任意のコマンド→DUMMY1~4に割当	DBI
15	ECHANGE	内部演算記号→個別DBの演算記号	DBI
16	NCHANGE	個別DBの名前→内部名に変更	DBI
17	TCHANGE	コード → 任意の文字列	DBI
18	OCHANGE	任意出力フォーマットの追加	DBI