九州大学学術情報リポジトリ
Kyushu University Institutional Repository

# SIGMAによる公用データベース昆虫学データベース(KONCHU)の公開とその利用法

多田内, 修
九州大学農学部昆虫学教室

https://doi.org/10.15017/1468166

出版情報：九州大学大型計算機センター広報. 20 (6), pp.597-614, 1987-11-25. 九州大学大型計算機センター
バージョン：
権利関係：

# （ ＳＩＧＭＡによる公用データベース 昆虫学データベース（ＫＯＮＣＨＵ） の公開とその利用法

多田内　修*

## １．はじめに

科学技術庁編の「データベースの高度利用」[10]で指摘されているように、生物に関するデータは、データベースの格好な対象となるべき特質を備えている．なぜならば、生物は有史以来その形態や機能が記載され続けてきており、また生物学とその周辺領域の発展とともに、データとして記載される内容は、絶えず追加・削除・変更されてきているからである．この蓄積された膨大な情報を有効に利用するために、生物データのデータベース構築を核とする情報システム開発への要求が高まっている．

筆者は、1928年に創刊された福岡蟲の会発行の昆虫学雑誌「ＭＵＳＨＩ（蟲）」全５０巻(1928年～1985年)の索引誌を、1985年に作成した[16]．これは ＭＵＳＨＩ に掲載されている昆虫の分類単位（種名、属名、科名などの学名または和名）をレコードとした、書誌情報とキーワードからなっている．このデータを九大理学部基礎情報学研究施設で開発されたテキストデータベース管理システム ＳＩＧＭＡ [2] によりデータベース化し、九大大型計算機センターの公用データベース [12] として登録した．これを契機として、その後「昆虫学データベース(ＫＯＮＣＨＵ）」の構築を計画し、実行に移している．これは昆虫学に関する1900年以後の国内の主要昆虫学雑誌をもとにしたデータベースで、各文献中の昆虫の分類単位を一つのレコードとしている．ＭＵＳＨＩ の索引誌作成時より、データの項目数を大幅に増やし、単なる文献データベースではなく、昆虫の様々な属性をデータとして含む新しいファクトデータベースとしての性格を持っている．今回データベース管理システムであるＳＩＧＭＡ 第２版([3]、[4]、[5] ）が九大大型計算機センターで公開されることになった．それに合わせて、現在構築の完了している２つのファイル「ＭＵＳＨＩ」と「ＥＳＡＫＩＡ」を同時に公開するので、ここにその利用法と実行例を紹介する．現在文部省科学研究費「研究成果公開促進費」により、日本昆虫学会誌「Ｋｏｎｔｙｕ̂（昆蟲）」のデータベース化を進めており、今後日本の主要な昆虫学雑誌を１年１～２誌の計画で同様な方法で作成していく．構築の完了したファイルは、昆虫学データベースのファイルの一部として順次公開の予定である．また、本解説文の他にシソーラス（検索語集）を印刷準備中である．

## ２．昆虫学とデータベース

生物関連のデータベースとしては、生物の様々な属性をデータとして含むファクトデータベースと文献データベースの２つが考えられる． 後者の場合には１文献を１レコードとし、書誌データから

昭和62年9月30日受理

* 九州大学農学部昆虫学教室

構成される（抄録を含む場合と含まない場合がある）という形式がほぼ確立している．一方、前者の場合にはさらにデータの要素により文字、数値、画像の３種類に分けられる．ファクトデータベースをもとにして、生物の同定をする試みは、Pankhurst [14] 以来研究が進められ、筆者も、文字と数値のデータファイルをもとにして、同定、記載、形態解析、分類群比較、分類群作成など一連の分類作業のできる総合的分類作業システム ＳＡＣ を開発し、ヒメハナバチ(姫花蜂)というハチのグループの分類学的研究に応用した[15]．ファクトデータベースを構築する場合、地球上の何百万種もの生物を整理する一つの方法は、国際命名規約（動物、植物、細菌の３種類がある）に従ってつけられた生物の学名を、生物に関するデータベースのもっとも基本的な単位とすることである． 今回公開する「昆虫学データベース」は各文献中に掲載されている種や属といった学名（または和名）を一つのレコードの単位とし、レコードごとに書誌情報と様々な情報を再編集した新しいファクトデータベースの性格を持っている．例えば、新種の場合には、分類学、新種の表示とともに、記載されている雌雄の別、記載のもとになった模式標本（ホロタイプ）の雌雄の別、模式標本の保存機関と標本の整理番号、模式標本の産地(模式産地)、分布、その他生態的属性などが含まれる．そのためにはデータ作成の際に論文にすべて目を通し，すべての学名を抜き出し、さらに項目ごとに必要なデータを選び出すという作業が必要であり、内容を充分理解把握した上で編集を行わねばならない．.

学名をデータベースの基本的単位とする場合に問題がないわけではない．ある昆虫に用いられる学名は固定的または安定したものではない．既存の種や群の再検討により同物異名（synonym）が発見されて学名が訂正されたり、属の新設や所属する属の変更に伴う学名の変更は枚挙にいとまがない．また、特に高次分類階級については、研究者により異なる分類体系が用いられる場合がしばしばあり、異なる名称が併用されることも多い．こうした学名の変更や異なる分類体系間の「翻訳」の問題については、後述する ＢＩＯＳＩＳ と Zoological Record が、1975年以来 ＴＲＦ（International Taxonomic Reference File）と呼ぶファイルの研究プロジェクトを開始し、テストファイルを使って研究を続けている（[6]、[7]）． 「昆虫学データベース」では、同物異名の項目（ＳＹＮ）を設けているので、学名の変更にも基本的に対応できるようになっている．

現在生物学関連の文献データベースとしてもっともよく知られているものは、アメリカのフィラデルフィアにある ＢＩＯＳＩＳ（BioSciences Information Service）である．1926年以来活動を続け、抄録誌 Biological Abstracts を発行しており、1969年度分よりデータベース化（初期のデータベースには抄録はない）が行われ、年間収録件数は２つのファイルＢＡ，ＢＡ／ＲＲＭ を合わせて現在では約３５万件を数える．大阪大学大型計算機センターには，最近ＢＩＯＳＩＳが導入され[17]、サービスが開始されている．また、動物学関連では ロンドン動物学会によって1864年に開始され、1886年以来100年以上発行が続けられている Zoological Record という索引誌がある． 1980年、ロンドン動物学会とＢＩＯＳＩＳとの交渉に基づいて、翌1981年から1978年度分の第115巻を両者が共同発行するようになった． その結果データは機械可読型となりデータベース化され、最新の情報は速やかにオンラインで入手できるようになった．

昆虫類は世界で約100万種、日本国内だけでも2～2.5万種は生息すると考えられている．毎年膨大な数の文献が発生しているが、特に昆虫学においては分類学のみならず、応用昆虫学の分野において

も過去にさかのぼって文献検索を行わねばならない．昆虫の種名の決定に際して50年以上前の文献を調べることは昆虫学研究者にとっては日常的なことである．最新の情報は ＢＩＯＳＩＳ 等のデータベースで検索が可能であるが、前記のように20年以上前の文献のレコードは国際的にもまだ存在しないと思われる．また、和文で書かれた国内文献の大半は、このようなデータベースには収録されていない．1800年代以前から蓄積のある昆虫学部門ではＢＩＯＳＩＳ の利用価値は限定されている．これを補完する意味で、当面国内文献のデータベース化に限ってみても日本及び東アジアの昆虫学を推進する上で意義あるものと思われる．今回の公開により、大学間ネットワークシステムを通じて全国の研究者に有効に活用していただければ幸いである．

## ３．昆虫学データベースの構成

昆虫学データベースは国内主要昆虫学雑誌の各々を一つのファイルとして構成されている．現在構築の完了しているファイルは次の２件である．

### １）ＭＵＳＨＩ

雑誌名「ＭＵＳＨＩ（蟲）」．福岡蟲の会発行で、1928年に創刊され、国際的昆虫学雑誌に発展し、1985年発行の第50巻をもって終刊した．50巻の通常巻の他に Supplement が３巻ある．ＭＵＳＨＩの歴史については中尾の解説があり[13]、第50巻には著者名索引と分類名索引 [16] がある．「ＭＵ－ＳＨＩ」ファイルのレコード数は、8,870件である．データベース検索の際のファイル名は、S.'A71414B. MUSHI'（著作権者：多田内修）である．

### ２）ＥＳＡＫＩＡ

雑誌名 ＥＳＡＫＩＡ (Kyushu University Publications in Entomology). 1960年に九州大学農学部附属彦山生物学実験所の昆虫学論文集として第1号が発行され、その後、1978年の第11号より九州大学農学部昆虫学教室と共同の編集発行となり、九州大学農学部昆虫学論文集となっている．現在までに25号（1987年2月発行）が発行されている．「ＥＳＡＫＩＡ」ファイルのレコード数は、3,063件である．今後新しく発行されるごとにデータを更新していく．データベース検索の際のファイル名は、S.'A71414B. ESAKIA'（著作権者：多田内 修）である．

尚、現在日本昆虫学会誌「Ｋｏｎｔｙｕ̂」のデータを作成中で、ファイルが完成次第昆虫学データベースのファイルの一部に加える．

## ４．テキストデータベース管理システム　ＳＩＧＭＡ 第２版

昆虫学データベースのデータベース管理システム（ＤＢＭＳ）は、今回公開されるテキストデータベース管理システム ＳＩＧＭＡ 第２版である．システムの詳細とその使用例については、本稿とともに広報 [5] に掲載されるので参照されたい．ここでは昆虫学データベースを利用する上で必要なシステムの知識と概要を、文献（[2]、[3]、[4]）に従って紹介する．

このシステムは、1981年に研究者向きの情報システムとして公開された[2]．これは文献の蓄積・

検索、論文の作成、自然言語の解析、書類整理、その他データの収集・加工など、研究者の日常的な活動を支援する目的で開発された．今回の第２版（[3]、[4]）では、日本語テキストが使用可能となり、個人用に使用可能な ＭＥＭＯ 領域が大幅に拡大され、ソート機能が強化されている．日本語が使用可能になったことで昆虫学データベースでの和文論文、和名その他和文データへの対応ができるようになり、たいへん使いやすくなったといえる．ＳＩＧＭＡ はファイル全体を一度だけ先頭から一字ずつ走査する一方向逐字処理に基本を置いている．そこで、文献データベースの作成や検索、自然言語の解析 [9] などに活用されている．

１）ファイル格納の領域

ＳＩＧＭＡ はファイルの高速処理及び共有データベース管理のため独自のファイルシステムを持ち、ファイルを管理している．ファイルは用途ごとに同じ種類のものがまとめられ、領域と呼ばれるところに保管されている．領域には、データベースを共有する ＳＩＧＭＡ 領域、個人用の ＭＥ-ＭＯ 領域、外部データセットの集まりである外部領域の３領域があり、各領域間でファイルを転送できる．

２）部分領域（区域）

ＭＥＭＯ 領域はその内部がさらに次の５種類の部分領域（区域）に分かれる．

ａ．ＭＥＭＯ区域：検索の結果などを、個人用に整理したいときに、ここにファイル名をつけて保存しておく．名前付きのファイルを MEMO ファイルと呼ぶ．

ｂ．作業区域：検索、置き換え、ソートなどの作業の中間結果を置くための区域で、スタック構造をしている．この区域のトップ（W.1）にそのファイルが作られる．W.1 は単に W のみでもよい．ファイルには新しいものから順に W.1, W.2, W.3 ... W.20 のようなファイル名がついている．この他に、ＬＯＧ区域、索引区域、バックアップ区域もスタックになっている．

ｃ．ＬＯＧ区域：使用したコマンド類はシステムとの交信記録として、この区域に自動的に保存される． ファイル名は新しいものから順に L.1, L.2, L.3 ... である．

ｄ．索引区域：検索コマンドの索引付け用のファイルを保存する．ファイル名は I.1, I.2, I.3 .. である．

ｅ．バックアップ区域：MEMO ファイルを除去した時とオーバーライト（上書き）した時に、この区域に移される．また、各スタックのファイルが20個を越えた時、最古のファイルがこの区域に自動的に移される．この区域のファイルが900個になると最古のファイルから順に完全に消去される．このようにいわば屑篭のように、容量に余裕が有る限り消されたファイルもシステム上では保存され、操作ミスによるデータの損失を防いでいる．

３）ＳＩＧＭＡ ファイルの共有

昆虫学データベースで検索の対象となる各ファイルはＳＩＧＭＡ 領域に格納されており、この領域のファイルはすべて先頭に S. を持つ（例えば、S.▼A71414B.ESAKIA▼）． 図１．は昆虫学データベ

ースを利用する際の、利用者1（データベース成作者）と利用者2（一般利用者:リード・アクセスのみ）のSIGMA領域とMEMO領域を示している．利用者はSIGMA領域を共有し、個人のMEMO領域の作業区域のファイルを用いて、検索、置き換え、ソートなどを行い、個人用に整理したければ名前を付けてMEMOファイルとする．

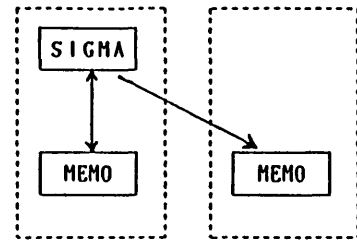


図1．SIGMAファイルの共有．

4）コマンド

SIGMAでのコマンドとその機能を表1．に示す．コマンドの入力には、下線部の最短省略形でもよい．

表1．SIGMAのコマンド一覧

| | |
|---|---|
| CATENATE | ：ファイルを連接する． |
| COMMENT | ：ファイルにコメントをつける． |
| COPY | ：ファイルの複製を作る． |
| DDIRECTORY | ：ディレクトリ（ファイル名一覧）を詳細情報とともに表示する． |
| DELETE | ：ファイルを除去して（消去ではない）、バックアップ区域のトップに置く． |
| DIRECTORY | ：ディレクトリを表示する． |
| ECHO | ：ログファイルをコマンドプロシジャとして実行中に、実行内容を表示するかどうかを設定したり、端末にメッセージを出力したりする． |
| END | ：SIGMAシステムを終了する、DO状態を終える． |
| GET | ：指定するファイルを作業区域のトップに置く． |
| HELP | ：SIGMAで使用できるコマンドや用語の説明を表示する |
| HEXDUMP | ：ファイルの16進ダンプを取る． |
| KEYIN | ：作業区域のトップにあるファイルに、キーボードから直接入力する． |
| LIST | ：ファイルの内容を表示する． |
| LOAD | ：指定するファイルの内容を、作業区域のトップにあるファイルに読み込む． |
| LOOK | ：作業区域のトップにあるファイルの内容を表示する． |
| MOVE | ：ファイルを移動する |
| NICKNAME | ：ファイルに別名をつける． |
| PASSNUMBER | ：パスナンバーを設定したり、変更したりする． |
| PROFILE | ：システム定数を表示したり設定したりする． |
| PUT | ：作業区域のトップにあるファイルを指定する所に置く． |
| REFILE | ：検索結果を再ファイル化する． |
| REPLACE | ：複数文字列の同時置き換えをする． |
| SAVE | ：作業用区域のトップにあるファイルの内容を指定するファイルへ書き出す． |
| SEARCH | ：複数キーの同時逐字検索を行う． |
| SETMEMO | ：MEMO領域を初期化する． |
| SETSIGMA | ：SIGMA領域を初期化する． |
| SORT | ：レコードのソーティングをする． |
| SPACE | ：MEMO領域の使用可能量を表示する． |
| TSS | ：TSSコマンドを呼び出す． |

5．昆虫学データベースのデータの形式

昆虫学データベースで扱うレコードとは、1論文中に掲載されている分類学上の1分類単位の学名（Taxon: 種、属、科など）である．従って、同一論文中で離れたページに同じ分類単位が出てきた場合には、一つのレコードとして扱う．それ故、別論文に出てくる同一分類単位は別のレコードとなる．雑誌「ＭＵＳＨＩ」の初期の和文論文では学名がなく和名のみの場合があり、その際には和名をレコード単位として採用した．レコード単位の大部分は種名であるが、科や属などのモノグラフ、再検討などの論文からは、科名、属名、時に目名などもレコード単位として採用した．クモ類、ダニ類は昆虫類には含まれないが、各文献中に掲載されている関係上便宜的に昆虫と同様な方法でデータを作成している．

データの項目とタグの説明は下記の通りである．

1．分類単位名（学名）（ＴＡＸ）
2．分類単位名（和名）（ＪＴＡＸ）
3．著者名（ＡＵ）
4．論文タイトル名（Ｔ）
5．雑誌名（Ｊ）
6．巻・号・頁（ＶＮＰ）
7．刊行年（Ｙ）
8．目名（ＯＲ）
9．科名（ＦＡＭ）
10．同物異名（ＳＹＮ）
11．分布（ＤＳＴ）
12．キーワード（ＫＥＹ）　別にシソーラスを印刷準備中である．

実例を下に示す．

```
#
(TAX) Inostemma seoulis (Ko), n. comb.
(JTAX)マツタマヤドリクロコバチ
(AU)  YOSHIDA, Nariaki;HIRASHIMA, Yoshihiro
(T)   Systematic studies on Proctotrupoid and Chalcidoid parasites of gall midge
s injurious to Pinus and Cryptomeria in Japan and Korea (Hymenoptera)
(J)   Esakia
(VNP) (14): 115-118
(Y)   1979
(OR)  Hymenoptera
(FAM) Proctotrupoidea, Platygasteridae
(SYN) Isostasius seoulis Ko
(DST) Japan(Hokkaido, Honshu, Kyushu);Korea
(KEY) systematics;redescription(female, male);new combination;parasite of Thecod
iplosis japonensis;マツバノタマバエ;biology;gall midges;Pinus
#
```

各レコードについて１２個の項目データがすべて論文中にあるとは限らず、和名や同物異名、分布などのデータがない場合には、項目は設けていない．論文中の synonymic list に記載されている学名はすべて同物異名の項目（ＳＹＮ）に入力してあるので、旧名でも検索が可能である．

データ作成上、文字の扱いに関していくつか問題が生じたが下記のように入力してあるので、検索する際には注意が必要である．

ａ．アルファベットは大文字、小文字の区別をしている．従って大文字は一般に使用されているように、固有名詞の第１文字としてのみ用い、例外的に著者名項目（ＡＵ）の姓名部分のみすべて大文字でデータを作成している．学名の属名の第１文字は大文字で始まり、種名は小文字で始まる．

ｂ．英数字は１バイト文字、ひらがな・カタカナ・漢字は２バイト文字でデータを作成している．和名などカタカナで検索する場合には、２バイト文字で入力する必要がある．

ｃ．ＳＩＧＭＡ システムで扱わない下記の記号は次のように置き換えているので、検索時には注意が必要である．

á→ 'a

â→ ^a

ä→ "a

à→ ～a

å→ %a

ç→ *c

ｄ．空白（ブランク）も一つの文字として意味を持つ．キーワードの登録には注意しなければならない．

６．ファイルの検索

ＳＩＧＭＡ システムでのファイルの検索は、SEARCH コマンドを用いて行われる．このコマンドは、データを何ら加工することなく、そのままの形で前から後ろに一読し、その間にすべての検索処理を行う一方向逐字処理の方法を用いている．SEARCHは 複数の検索キーワードを用いた複数の質問を同時に処理できる特徴を持っている．

１）レコード区切り語

SEARCH コマンドの検索の単位は、レコード区切り語によってはさまれた部分で、レコードと呼ばれる．レコード区切り語として任意の文字列を用いることができるが、昆虫学データベースのファイルではレコード区切り語として "#" 記号を用いているので、システムからの質問には "#" を入力すればよい．

２）項目区切り語

検索する際の質問式の評価を行う範囲を指定するもので、項目区切り語を検出した時点で、質問式の評価が行われる．これによりきめの細かい検索を行うことができる．昆虫学データベースのフ

ァイルで項目を指定して検索を行う場合には、システムからの質問に "!" 記号を入力する．項目を指定しない検索の場合には入力の必要はない．

3）キーワード

レコード区切り語及び項目区切り語の指定を終わると、次にキーワードの登録をしなければならない．キーワードは任意の文字列である．SEARCH では同時にキーワードを複数個登録することができ、入力された順にキーワード変数（A1, A2, A3 ...)が付けられる．キーワードとしては、通常の文字列の他に、トリプル・ドット" ..." を含む文字列も使用できる．トリプル・ドットとは、任意の文字列（空でもよい）を表すもの（ワイルドカード）であり、" A ... B" は Aがあり、続いてある文字列の後にBがあることを示している．

4）論理式

キーワードの登録が終わると、直接キーワードを使わず質問のための論理式を入力する．論理式はキーワード変数（A1, A2 ...)、式変数（Z1, Z2 ...)、整定数、論理演算子（and (.), or (,), not (^), 括弧 ((,)) ）などを組み合わせて作る．代表的な組み合わせを下に示す．

| | |
|---|---|
| Ｚ１：＝Ａ１．Ａ２ | Ａ１かつＡ２ |
| Ｚ２：＝Ａ１，Ａ２ | Ａ１またはＡ２ |
| Ｚ３：＝＾Ａ１ | Ａ１ではない |

5）ファイルの指定

論理式の登録が終わると、検索結果を表示するかどうかをきいてくるので、必要ならば Y を選ぶ．最後に検索の対象となるファイル名を入力する．ＥＳＡＫＩＡ ファイルの場合には、正式には S. 'A71414B.ESAKIA' であるが、次の使用例で説明するように最初に PROFILE コマンドの PREFIX サブコマンドで、A71414B を指定しておけば、ファイル名は S.ESAKIA で省略できる．ＭＵＳＨＩ ファイルの場合には、S.MUSHI と指定する．これでシステムは検索を開始する．

6）検索結果の表示

検索が終了すると、見つかったレコード数と、検索に要した CPU 時間（単位はミリ秒）が表示され、さらに検索すべき別のファイルがあるかどうかをたずねてくる．ここでファイル名を入力せずに単に復改(リターン)キーを押せば DO 状態に戻る．プロンプト DO: のあとに REFILE コマンドを入力すると検索結果を表示するか否かをきいてくるので、Y か N を入力する．プロンプト QUESTION:= に従って希望する質問番号(論理式の番号)を入力し、さらに新しいレコード区切り語、検索されたレコードにナンバーをつけて表示するかいなか、出力ファイルを順に指示する．出力ファイルを画面（Dまたは T）にすると，結果が端末に表示される．プロンプトQUESTION:= の後に復改を入力すると、REFILE コマンドを終了する．

## 7．ESAKIA ファイルを用いた検索の具体例

以下に ESAKIA ファイルを用いた検索の具体例を、TSS のセッションの開始から終了までを含めて紹介する．アンダーラインの部分は利用者が入力する．

使用例 1．

```
     JCET005 SYSTEM READY
(1)  LOGON TSS A79999A
     + PASSWORD ?
(2)  Password
     KDS700011 A79999A  LAST ACCESS AT 17:25:43 ON 87.267
     JOB NO = TSU3857 CN(01)
     KEQ564551 A79999A LOGON IN PROGRESS AT 17:34:02 ON SEPTEMBER 24, 19
     87
     KEQ569511 NO BROADCAST MESSAGES
     READY
(3)  SIGMA
     THE CREATION OF SPACE IS STARTED.
     THE NUMBER OF BLOCKS MUST BE FROM 10  TO 4776 .
(4)  NUMBER OF BLOCKS:=50
     NUMBER OF BLOCKS      =   50
     NUMBER OF WORK   FILES=   20
     NUMBER OF LOG    FILES=   20
     NUMBER OF INDEX  FILES=   20
     NUMBER OF BACKUP FILES=  900
(5)  MASTER KEY:=1111
       1144000 BYTES ARE AVAILABLE.
     THE CREATION PROCESS IS COMPLETED.
(6)  SIGMA>PROFILE
(7)  PROF:DISPLAY
     ****** SYSTEM CONSTANTS ******
     NEWLINE CHARACTER =    !
     CONTINUE CHARACTER =   -
     TRIPLE DOT CHARACTER = .
     LINE SIZE =             78
     BELL MODE =            ON
     PREFIX OF MEMO =
     PREFIX OF SIGMA =      A79999A
     SYSOUT CLASS =         U
     HELP LEVEL =           0
(8)  PROF: PREFIX
     MEMO OR SIGMA? (M OR S) S
(9)  PREFIX:=A71414B
(10) PROF:DISPLAY
     ****** SYSTEM CONSTANTS ******
     NEWLINE CHARACTER =    !
```

```
     CONTINUE CHARACTER =   -
     TRIPLE DOT CHARACTER = .
     LINE SIZE =             78
     BELL MODE =            ON
     PREFIX OF MEMO =
     PREFIX OF SIGMA        A71414B
     SYSOUT CLASS =         U
     HELP LEVEL =           0
(11) PROF:SAVE
     SYSTEM CONSTANTS SAVED
(12) PROF:END
(13) DO:DIRECTORY S
     PASSNUMBER:=↙
     ESAKIA
     MUSHI
              TOTAL =   2
      2624 SECTOR(S) AVAILABLE
(14) DO:SEARCH
          RECORD DELIMITERS
(15) DO1:=#
     DO2:=↙
          ITEM DELIMITERS
     DO2:=↙
          KEYWORDS
(16) AO1:=Noxius
     AO2:=↙
          LOGICAL FORMULAS
(17) ZO1:=A1
     ZO2:=↙
(18) REPORT (Y/N)?Y
(19) FILE:=S.ESAKIA
        RETRIEVED TEXTS
     QUESTION 01 (ZO1) =        2         2
     TOTAL             =        2         2
     CPU (SEC/1000)    =      344       344
(20) FILE:=↙
(21) DO:REFILE
(22) REPORT (Y/N)?Y
        RETRIEVED TEXTS
     QUESTION 01 (ZO1) =        2
     TOTAL             =        2
(23) QUESTION:=1
(24) NEW RECORD DELIMITER:=#
(25) NUMBERING (N/Y)?↙
(26) OUTPUT-FILE:=T
     QUESTION 01 (ZO1) =        2
```

```
     #
     (TAX) Noxius Jordan
     (AU)  MORIMOTO, Katsura
     (T)   The family Anthribidae of Japan (Coleoptera). Part 4
     (J)   Esakia
     (VNP) (17): 54-56
     (Y)   1981
     (OR)  Coleoptera
     (FAM) Anthribidae
     (SYN) Blaberus Schoenherr
     (DST) Japan
     (KEY) systematics;revision;include 1 new sp.
     #
     (TAX) Noxius japonicus Morimoto, n. sp.
     (AU)  MORIMOTO, Katsura
     (T)   The family Anthribidae of Japan (Coleoptera). Part 4
     (J)   Esakia
     (VNP) (17): 54-56
     (Y)   1981
     (OR)  Coleoptera
     (FAM) Anthribidae
     (DST) Japan(Kyushu, Ishigaki Is., Yonaguni Is.)
     (KEY) systematics;new species;description(male, female);type(male:K
     yushu Univ., No. 2296);type locality(Mt. Urabu, Yonaguni Is.)
     #
(27) QUESTION:=↙
(28) DO:END
(29) SIGMA>END
     READY
(30) LOGOFF
     RETURN CODE : 0000
     KEQ564701 A79999A LOGGED OFF AT 17:37:06 ON SEPTEMBER 24, 1987+
     KEQ542201 SESSION ENDED
```

1） 使用例1の説明（起動と終了、検索）

（1）ＴＳＳセッション開始のためのコマンドとユーザーＩＤ を入力する．

（2）パスワードを入力する

（3）ＳＩＧＭＡを起動するには、READYモードで " SIGMA" と入力する．

（4）利用者が初めて ＳＩＧＭＡ を使う場合には、個人用の MEMO ファイルのスペースを確保しておかなければならない．ブロック数は50で約１メガバイトになるので、利用目的に沿って適当な数値を入力する．

（5）個人用の MEMO ファイルには個人専用のマスターキーを使うことができる．マスターキーは領域の初期化などに必要となってくるので、記憶されたい．４桁の数字からなり、単に復改キーを押すと、0となる．

（６）" SIGMA>" は、ＳＩＧＭＡ のプロンプトで、この状態（SIGMA 状態）の時コマンドを入力できる．システム定数の変更を行うために、コマンド PROFILE を入力する．

（７）プロンプト" PROF:" に続いてサブコマンド DISPLAY を入力し、システム定数の現在値を表示する．

（８）ＳＩＧＭＡ 領域にあるファイルを利用するには、ここでＳＩＧＭＡ 領域 のプレフィックスを変更しておくと便利である．サブコマンド PREFIX を入力し、S を指定する．初期値と省略値はユーザーＩＤである．

（９）「昆虫学データベース」を利用するには、ＳＩＧＭＡ 領域にあるユーザーＩＤを指定すると便利である．そのためにプレフィックスとして、ユーザーＩＤである "A71414B" を入力する．

（10）変更したシステム定数を確認するために、サブコマンド DISPLAY を入力する．

（11）システム定数の現在値を保存する．これを行わないと、ＳＩＧＭＡ から出た時点で変更値は失われる．

（12）サブコマンド END を入力し、PROFILE を終わる．復改でもよい．

（13）" DO:" は、（６）ででた " SIGMA>" と同じＳＩＧＭＡのプロンプトと考えてよい．" DO:" が出されている状態（DO状態）でコマンドを入力することができる．DIRECTORY コマンドによりＳＩＧＭＡ領域のファイル名一覧を表示する．"PASSNUMBER:=" のプロンプトに対しては、復改のみを入力すればよい．これにより２つのファイルがあることがわかる．

（14）SEARCH コマンドにより検索を行う．コマンドの入力には、最短省略形までのどの形でもよい．SEARCH の場合には、SEA、SEAR、SEARC、SEARCH のいずれでもよい．

（15）レコード区切り語 "#" を入力する．次の項目区切り語は項目を指定して検索するのでなければ、必要ないので復改のみを入力する．

（16）キーワードを、キーワード変数 "A01:=" に続いて入力する．ここでは "Noxius" という甲虫の属名を入力している．キーワードは同時に複数個検索できるので、その際には次のキーワードを " A02:=" に続け、さらに検索する場合には同様に順次キーワードを入力していく． 登録を終わるには、復改のみ入力する．

（17）キーワードを登録したキーワード変数により質問を作り、" Z01:=" に続けて入力する．使用例では、A1（A01のかわりにA1でよい）というもっとも単純な質問をしている．登録を終わるには復改のみを入力する．

（18）検索結果を表示するか否かを入力する．省略値は "Y" である．"Y" を入力している．

（19）"FILE:=" に続いて、検索するファイル名を入力する．ＥＳＡＫＩＡ ファイルの場合には、"S.ESAKIA" と入力する．その結果、検索個数とＣＰＵ時間が表示される．ここでは２個見つかっている．

（20）同じ質問を別のファイルで行う場合には、ここでそのファイル名を入力する．ここでは行わないので、復改キーを押している．

（21）コマンド REFILE を入力する．検索結果をさらに再ファイル化するためのコマンドである．

（22）SEARCH の検索結果を再度表示するか否かをきいてくる．ここでは、"Y" を入力している．

(23) 質問番号を入力する．論理式の "Z01" を質問する場合には、"1" と入力する．

(24) 新しいレコード区切り語を、もとのまま "#" として指定する．

(25) 検索結果にナンバーをつけるか否かをきいてくる．復改の場合は、"N" となる．

(26) "OUTPUT-FILE:=" に続いて、"T" と入力して、検索結果を端末に出力させる．質問番号に対する検索個数と検索されたレコードが表示される．

(27) 質問を終わる場合には、"QUESTION:=" に続けて復改を押す．

(28) DO状態を終わるコマンド END を入力している．

(29) SIGMA を終わるコマンド END を入力している．READYモードに戻る．

(30) セッションを終わる．

使用例 2．

```
(1)  DO:SEA
          RECORD DELIMITERS
     DO1:=#
     DO2:=↙
          ITEM DELIMITERS
     DO2:=↙
          KEYWORDS
(2)  A01:=n. sp.
     A02:=n. subsp.
     A03:=n. gen.
     A04:=n. subfam.
     A05:=n. syn.
     A06:=↙
          LOGICAL FORMULAS
     Z01:=A1
     Z02:=A2
     Z03:=A3
     Z04:=A4
     Z05:=A5
     Z06:=↙
     REPORT (Y/N)?↙
     FILE:=S.ESAKIA
        RETRIEVED TEXTS
(3)  QUESTION 01 (Z01) =        516        516
     QUESTION 02 (Z02) =         25         25
     QUESTION 03 (Z03) =         13         13
     QUESTION 04 (Z04) =          1          1
     QUESTION 05 (Z05) =        135        135
     TOTAL             =        690        690
     CPU (SEC/1000)    =        371        371
     FILE:=↙
     DO:
```

2）使用例2の説明（検索）

（1）SEARCH コマンドを入力する．

（2）キーワードを入力する．ここでは、新種（n. sp.: new species でもよい．以下同様)、新亜種（n. subsp.)、新属（n. gen.)、新亜科（n. subfam.)、新シノニム（n. syn.）を キーワードとして登録している．

（3）ＥＳＡＫＩＡファイルに含まれるそれぞれの個数が、516、25、13、1、135 であることがわかる．

使用例 3．

```
(1) DO:SEARCH
          RECORD DELIMITERS
    DO1:=#
    DO2:=↙
          ITEM DELIMITERS
    DO2:=!
    DO3:=↙
          KEYWORDS
(2) AO1:=(TAX) Palaeorhiza
    AO2:=(DST)...New Guinea
    AO3:=(DST)...Australia
    AO4:=Yakushima
    AO5:=Amami-Oshima
    AO6:=シェーファー
    AO7:=↙
          LOGICAL FORMULAS
(3) ZO1:=A1
    ZO2:=A2
    ZO3:=A3
    ZO4:=A4
    ZO5:=A5
    ZO6:=A6
    ZO7:=Z1.Z2
    ZO8:=Z1.Z3
    ZO9:=Z4.Z5
    Z10:=↙
    REPORT (Y/N)?↙
    FILE:=S.ESAKIA
       RETRIEVED TEXTS
(4) QUESTION 01 (ZO1) =        150        150
    QUESTION 02 (ZO2) =        399        399
    QUESTION 03 (ZO3) =         54         54
    QUESTION 04 (ZO4) =        115        115
    QUESTION 05 (ZO5) =        140        140
    QUESTION 06 (ZO6) =          1          1
    QUESTION 07 (ZO7) =        125        125
```

```
    QUESTION 08 (Z08) =        10        10
    QUESTION 09 (Z09) =        42        42
    TOTAL             =       665       665
    CPU (SEC/1000)    =       887       887

    FILE:=↙
    DO:REFILE
    REPORT (Y/N)?N
(5) QUESTION:=6
    NEW RECORD DELIMITER:=#
    NUMBERING (N/Y)?↙
    OUTPUT-FILE:=T
    #
    (TAX) Eurytoma schaeferi Yasumatsu et Kamijo, n. sp.
    (JTAX)シェーファーカタビロコバチ
    (AU)  YASUMATSU, Keiz^o;KAMIJO, Kazuaki
    (T)   Chalcidoid parasites of Dryocosmus kuriphilus Yasumatsu (Cynipidae) in
    Japan, with descriptions of five new species (Hymenoptera)
    (J)   Esakia
    (VNP) (14): 106-109
    (Y)   1979
    (OR)  Hymenoptera
    (FAM) Chalcidoidea
    (DST) Japan(Hokkaido, Honshu, Kyushu, Satsunan Is.)
    (KEY) systematics;new species;description(female, male);type(female:Kyushu U
    niv., No. 2144);type locality(Teineyama, Hokkaido);biology;parasite of Dryoc
    osmus kuriphilus;chestnut
    #
    QUESTION:=↙
    DO:
```

1）使用例3の説明（検索）

（1）SEARCH コマンドを入力する．

（2）キーワードとして、A01 では、項目のタグを含めて "(TAX) Palaeorhiza" を入力している．A02、A03 では、ワイルドカード(...) を使っている．(DST) の次に、ある文字（空列でもよい）をおいて New Guinea、Australia という文字列のあるものを指定している．A06 ではカタカナを キーワードとしている．この場合には2バイトで入力する．検索する際に、正確にキーワードを覚えていない時には、このように正確な部分のみを指定してもよい．

（3）質問式を作る．Z07 は Z01 かつ Z02 であるもので、この場合 A1.A2 ではなく、Z1.Z2 と指定する．この意味は "(TAX) Palaeorhiza" というデータがあり、かつ (DST) の項目に "New Guinea" というデータのあるものである．つまり、New Guinea 産の Palaeorhiza を検索している．

（4）各質問に対するヒットした件数が表示される．New Guinea 産の Palaeorhiza は125件あることがわかる．同様に Australia 産の Palaeorhiza は10件ある．

（5）質問番号として "6" を入力している．シェーファーとしか覚えていなかった和名がシェーファーカタビロコバチであることがわかる．

使用例 4.

```
(1) DO:SEARCH
          RECORD DELIMITERS
    DO1:=#
    DO2:=↙
          ITEM DELIMITERS
    DO2:=!
    DO3:=↙
          KEYWORDS
(2) AO1:=(OR)...Coleoptera
    AO2:=(AU)...CH^UJ^O
    AO3:=(FAM)...Tenebrionidae
    AO4:=↙
          LOGICAL FORMULAS
    ZO1:=A1
    ZO2:=A2
    ZO3:=A3
    ZO4:=↙
    REPORT (Y/N)?↙
    FILE:=S.ESAKIA
       RETRIEVED TEXTS
QUESTION 01 (ZO1) =      2190      2190
QUESTION 02 (ZO2) =       524       524
QUESTION 03 (ZO3) =       443       443
TOTAL             =      2202      2202
CPU (SEC/1000)    =      1018      1018
FILE:=↙
DO:
```

4）使用例4の説明（検索）

（1）SEARCH コマンドを入力する.

（2）特殊記号の入力には注意が必要である. "CHUJO" は、"CH^UJ^O" と入力する.

8. 大学間ネットワークシステムによるアクセス

他大学の大型計算機センター利用者が大学間ネットワークシステムを利用して九州大学大型計算機センターを利用する場合には、あらかじめ、九州大学大型計算機センターに利用登録をしておく必要がある. 九大大型計算機センターのホスト・コンピューターにアクセスするには、全国共同利用大型計算機センター オンライン・データベース利用ガイド（毎年１１月に発行されている）、または、各地区共同利用大型計算機センターで発行しているネットワークシステムに関する利用の手引き等を参照されたい.

謝　辞

「昆虫学データベース」の公開にあたって、ＳＩＧＭＡ 第２版でのデータベース構築に協力をいただいた、九大総合理工学研究科情報システム学専攻の宮原哲浩氏、川崎洋治氏はじめ同研究科の院生諸氏に厚く御礼申し上げる．また、データ作成にあたって日頃有益な助言をいただく九大農学部平嶋義宏教授、森本　桂助教授、公開にあたって多くの助言と協力をいただいた九大理学部基礎情報学研究施設の有川節夫教授、九工大情報工学部の篠原　武助教授、九大大型計算機センターの二村祥一データベース室長に感謝申し上げる．

参考文献

1．Allkin, R. & F. A. Bisby (ed.): Databases in systematics. The Systematics Association special volume No. 26. 329pp. Academic Press, London. (1984).

2．有川節夫，篠原　武，白石修二，玉越靖司：研究者向き情報システム ＳＩＧＭＡ について．九州大学大型計算機センター広報，Vol.14, No.4: 550-573. (1981).

3．有川，武谷，原口，篠原，大島，白石，酒井，宮原，山本，井上，川崎，湯浅：汎用テキストデータベースシステム ＳＩＧＭＡ の改訂について．九州大学大型計算機センター計算機科学研究報告，No.4: 41-50. (1987).

4．有川，武谷，篠原，大島，原口，白石，川崎，井上，湯浅，酒井，山本，宮原：テキストデータベース 管理システム ＳＩＧＭＡ について．情報処理学会研究報告，Vol.87, No.65, FI-6-4: 1-8. (1987).

5．有川，篠原，武谷，大島，原口，白石，酒井，宮原，山本，井上，川崎，湯浅：テキストデータベース管理システム ＳＩＧＭＡ 第２版について．九州大学大型計算機センター広報，Vol.20, No.6. (1987).

6．Dadd, M. N.: Standardisation of biological names. An investigation of methods for use by information services. Brit. Lib. Res. Devel. Dep. Rep., No.5347. (1976).

7．Dadd, M. N.: A machine-readable taxonomic reference file. International Workshop January 29th-31st, 1979. Brit. Lib. Res. Devel. Dep. Rep., No.5548. (1979).

8．学術情報センター：学術情報データベース基本調査報告書．54pp. (1987).

9．樋口忠治，篠原 武 ：公用データベース トーマス・マン・ファイル/ＳＩＧＭＡ の公開．九州大学大型計算機センター広報，Vol.16, No. 4: 379－393. (1983).

10．科学技術庁編：科学技術情報活動の現状と展望 第７巻．データベースの高度利用．241pp. 大蔵省印刷局，東京．(1985).

11．九州大学大型計算機センター：利用の手引，ネットワーク編（第４版）．38pp. (1987).

12．松尾文碩，二村祥一，高木利久：公用データベースについて．九州大学大型計算機センター広報，Vol.15, No. 2: 222-227. (1982).

13．中尾舜一：「むし」５０巻の歩み－福岡虫の会小史－．Mushi, Vol.50, Part 6: 115-137. (1985).

14. Pankhurst, R. J.: Automated identification in systematics. Taxon, Vol.23: 45-51. (1974).

15. Tadauchi, O.: Taxonomic working system by computer (SAC) with application to Japanese Andrenid bees. ESAKIA, No.17: 161-182. (1981).

16. Tadauchi, O.: General index to Mushi Vol. 1(1928)-Vol. 50(1985). Subject index. A list from a database DMUSHI produced in the Computer Center of Kyushu University. Mushi, Vol.50: 47-175. (1985).

17. 多喜正城 :ＢＩＯＳＩＳ データベース・システム. 九州大学大型計算機センター広報, Vol. 20, No. 1: 1−24. (1987).

18. 全国共同利用大型計算機センターデータベース連絡会編：全国共同利用大型計算機センターオンライン・データベース利用ガイド第６版. 153pp. (1987).