

平和科学研究センター文献検索システム
PRINSE 日本語版

松 尾 雅 嗣

広島大学平和科学研究センター

**The Japanese Version of PRINSE, a Simple Retrieval
System for Peace Research Documents**

Masatsugu MATSUO

Institute for Peace Science, Hiroshima University

SUMMARY

This is a sequel to the former reports in this journal on PRINSE, a simplified peace research document retrieval system on the personal computer. The present one reports on its important latest improvements. The primary purpose of the improvement is to enable the program to deal with Japanese documents, which consist (mainly) of so-called 2-byte characters. In addition, many other additions and improvements are made, including such facilities as execution of user-defined macros, on-line manual, user customization, use of tags and so on.

はじめに

本稿は、平和科学研究センターの簡易文献検索システム PRINSE (松尾1985, 同1986, 松尾・竹花1986) の日本語文献対応版の報告である。この日本語版は、現在、平和科学研究センターの蔵書検索のために使用を開始している。今回の拡充の主眼は、日本語文献の入力、検索を可能にすることにあるが、このほかにも、プログラムの統合、マクロ等機能の追加を行なった。また、ユーザ・インターフェイスの拡充、就中操作性と利用の自由度の拡充も行なった。拡充、追加した主な機能は以下のとおりである。

日本語文献処理

自動実行、検索マクロ実行

ユーザカスタマイズ

タグの設定とタグによる検索

データベース切替

コマンドとメニューの切替

マニュアルのオンラインマニュアル化

1 実行

1.1 機器構成

PRINSE の実行に必要な機器、システム構成は以下のとおりである。

本 体：NEC PC 9801シリーズ

メモ リ：640K バイト以上

OS : MS-DOS V 3.1以上

記憶媒体：ハードディスクまたはフロッピーディスク

CRT : 必須、カラーが望ましい

プリンタ：あることが望ましい

また、日本語文献の処理のためには、日本語フロントエンドプロセッサが必要である。

1. 2 コマンドラインオプション

PRINSE は MS-DOS のコマンドとして実行される。実際に使用する時には、バッチファイルを作成しておいてそれを実行するのが便利であるが、これについてはここでは触れない。PRINSE のコマンドラインには、以下のコマンドラインオプションを指定することができる。

PRINSE 環境ファイル名

PRINSE メインコマンド

PRINSE 自動実行ファイル名

PRINSE データベースのパス名

このうち、データベースのパス名に関しては、2の「文献データとデータベース」で述べる。

1. 3 PRINSE の環境

PRINSE では、利用者がコマンドの名称、キーの操作などを変更できる機能、所謂ユーザカスタマイズ機能を実現するために、環境という概念を用いる。利用者は、環境を変更をすることにより、カスタマイズを行なう。PRINSE の環境は、環境ファイルと称する特別のファイルに記憶され、PRINSE は、起動時にこのファイルを読み込んでキーの割当て、コマンド名称、様々な実行用パラメータの値を得る。コマンドラインに環境ファイルの指定があれば、標準の環境ファイルではなく、指定された環境ファイルから必要な情報を得る。

環境の変更のためには、ふたつの方法が用意されている。ひとつは環境を一括して変更する方法であり、他のひとつは検索、文献データの入力中に変更する方法である。

前者は、利用者に提供される環境定義ファイルというファイルを書き換えるこ

とにより行なう。環境定義ファイルは、ユーザが変更可能なパラメータ（以下、環境変数と称する）の値をすべて標準値に設定した MS-DOS のテキストファイルであり、この内容を書き換えて、メインコマンドの「環境設定」機能を使用して環境ファイルを書き換える。環境定義ファイルには、すべての環境変数、即ちパラメータについて、簡潔な説明が注釈行として付されており、変更はごく容易である。

後者は、PRINSE 実行中に、環境のごく一部、例えばひとつの環境変数の値を変更することを目的とする。「環境の部分変更」あるいは「パラメータ変更」のメニューあるいはコマンドを利用して、変更を行なう。この変更の効果の持続範囲は、当該の PRINSE 実行中にのみ限定することも、次回に以降にも持続させることもできる。変更の効果の持続性自体も、ひとつの環境変数であり、実行中に変更、再設定できる。

1. 4 主たる機能とメインコマンド

PRINSE の主たる機能は以下のとおりであり、これがメインメニューとなる。メインメニューはそれぞれ PRINSE メインコマンド（以下、単にメインコマンドと呼ぶ）と 1 対 1 に対応している。

終了……………PRINSE の実行終了
文献検索……………文献の検索
新規作成……………文献データベースの初期値設定
データ入力／追加……………文献データの入力，追加
データベース情報……………作成されたデータベース情報の出力
データベース切替……………データベースの切替
タグ編集……………検索タグの設定，削除，表示
データ編集……………キーワード編集，文献の削除
索引／リスト出力……………キーワード，タグ，文献の出力，文献索引出力
環境の部分変更……………PRINSE 環境の実行中の部分的変更

環境設定……………PRINSE 環境の一括変更

マニュアル……………オンラインマニュアルの表示、印刷、ディスク出力

子プロセス実行……………子プロセスの起動と実行

補助機能……………PRINSE 旧版入力データの変換など。

このうち、「環境の部分変更」、「マニュアル」、「子プロセス実行」の各機能は、「検索」、「データ入力／追加」実行中にも使用できる。

メインコマンドの名称、メニューの名称は、ともに利用者が変更できる。他のサブコマンド名、(サブ)メニューの名称も同様である。

1. 5 実行モード

メインコマンドの実行については、3種類の実行モードがある。

ひとつは、コマンドラインオプションとして、メインコマンド名をコマンドラインに与える方式である。このときには、与えられたメインコマンド、例えば、データの入力／追加コマンド、をひとつだけ実行して終了する。このときには、次のメニューモードと異なり、メインメニューも表示されない。

第2は、前節に掲げた機能に対応するメインメニューを表示し、メニューから選択して実行する方式である。この方式では、選択されたメニューの実行が終了すると再びメニューの選択が可能となり、終了メニューが選択されるまで PRINSE の実行が継続される。これは、ここで述べるメインメニューのみならず、PRINSE のすべてのメニューに共通である。また、PRINSE のすべてのメニューでは、現在の多くのソフトが採用しているように、カーソルを移動して画面に表示されたメニューのひとつを選択する方法と、カーソル位置に関わりなく、メニューの各項目の先頭に示された1文字を入力して選択する方法のいずれもが同時に可能である。

第3は、コマンドラインにマクロ実行用ファイル名を与えて自動実行する方式であり、これについては、次節で述べる。

1. 6 自動実行マクロ

コマンドラインに自動実行ファイル名が与えられたときには、このファイルに与えられた処理を自動的に実行する。自動実行ファイル（もしくはマクロファイル）は、MS-DOS のテキストファイルであり、マクロ制御文、メインコマンド、検索用コマンドから構成される。

このうち、検索用コマンドは、検索を指示するメインコマンドが与えられた後にのみ使用可能である。

また、マクロ制御文は、主としてループによる反復処理の制御を目的とするものであり、以下のものが使用可能である。制御文の名称は、利用者が変更、設定できる。

ループの起点指定

ループの終端指定

ループからの脱出

キーボードからの入力

ファイルの取込（C 言語の #include に相当）

次の例は、検索を反復して行なうための、自動実行ファイルの例である。

find	①
%repeat	②
%input 検索コマンドは (%LEAVE=終了)	③
%x	④
%end	⑤
q	⑥
quit	⑦

①は、メインコマンドの「検索」コマンドである。これにより、文献の検索を

開始する。

②は、マクロ制御文のループ開始命令であり、⑤はループ終了命令である。この間にあるコマンドあるいはマクロ制御命令、つまり③と④が反復して実行される。この例からも明らかなようにこのループは無限ループであり、どこかでループ脱出のためのマクロ制御命令を与える必要がある。

③は、キーボードからの入力命令であり、この例のように、プロンプトを付けることができる。この行が実行されると、画面には、

```
検索コマンドは (%LEAVE=終了) =
```

が表示され、キーボードからの入力待ちとなる。入力は文字列として、マクロ変数、この場合は、%Xに代入される。④はこの変数に保存されたコマンドを実行する。従って、入力要求に対して、ループ脱出命令、標準では、%LEAVEを与えれば、ループの実行を終了して、⑥の検索終了コマンドが実行される。

この例では、検索実行中であるので、ループを脱出するとき以外には、入力要求に対して、後述の検索用コマンドのみを与えることができる。この例は、文献検索を検索を後述のコマンドモードで行なうのとはほぼ同じことを行なっている。

⑦は、PRINSEの終了を指示するコマンドである。

1. 7 エラー処理

実行時のエラーには3種類がある。

第1は、利用者の入力ミスなどの回復可能なエラーである。この場合は、画面にその旨のメッセージを表示して、再試行するかどうかを尋ねる。

第2は、メモリ不足など、修復不能のエラーである。この場合は、その旨のメッセージを表示して、実行を終了する。但し、他の処理が可能な場合には、上と同様の処理を行なうこともある。

第3は、所謂プログラムのバグであり、万一発生したときにはプログラムの暴走を引き起こす。

2 文献データとデータベース

2.1 文献データ

PRINSE の扱う文献データ 1 件は著者名、表題、書誌事項という 3 つのフィールドから成る。この他に、各文献に最大約 4 K バイトのコメントを付けることができる。コメントは、検索時にも参照、編集できる。

扱いうる文献データ、キーワード、タグ（後述）の最大件数は、理論上はすべて約 21 億件であり、実際には、メモリ、記憶媒体の容量の制約によって定まる。

2.2 PRINSE データベースとデータベースパス

PRINSE に入力された文献データは、索引、文献データ等目的に応じた数種類のファイルに分割して保存される。このファイルを総称して（PRINSE）データベースと言う。文献数、検索キーワード数など、データベースに関する情報は、メインコマンドの「データベース情報」により、出力できる。また、データベース中の文献、キーワード等のディスク出力、印刷に関しては、後の「補助的機能」を参照。

データベースは利用者、用途等に応じて幾つでも作成できるが、異なるデータベースは、別のディレクトリに（あるいは別のフロッピーディスクに）置くことを原則とする。データベースが置かれるディレクトリなどのパス名をデータベースパスと言う。複数の PRINSE データベースに同時にはアクセスできないが、「データベースパスの切替」機能を使用してデータベースを切替えて使用できる。

コマンドラインに、データベースパスを与えれば、起動時には、このパスにある PRINSE データベースが処理対象となる。複数のデータベースパスを、予め環境に登録しておくこともできる。

3 検索

3. 1 検索用コマンド

検索は指定された条件を満足する空でない文献の集合（以下単に文献集合と称する）を作成し、それを画面表示、印刷、ディスク出力するという手順で行なう。このための様々な指示は、画面に表示されるメニューを選択することによって（メニューモード）も、検索用のコマンドを直接投入することによって（コマンドモード）も可能である。また、自動実行時には、予め自動実行ファイルに書き込んでおくという方法も可能である。検索のために利用できるメニューは以下のとおりであり、これは検索用のコマンドと1対1に対応する。

検索キーによる検索

論理積論理和

著者順ソート

文献表示

文献外部出力

文献情報出力

文献情報入力

キー/タグ一覧

文献集合情報

検索マクロの定義、表示、削除

パラメータ変更

子プロセス実行

マニュアル参照

検索終了

3. 2 文献検索のパラメータ

文献集合を作成する最も基本的な方法は、言うまでもなく、特定のキーを有する文献の集合を作成することである。

利用者が与える検索のために与えるキー、即ち検索キーに関しては、アルファベットなど所謂アスキー文字に関しては、全角半角、大文字小文字の区別をしない。即ち、全角、半角いずれでも、あるいは大文字小文字いずれで入力してもよい。

検索キーに関しては、レベルの異なる2種類のキーを使用することができる。ひとつは、キーワードであり、これはデータベース作成時に、検索のキーとするよう指示された語句（正確には文字列）である。他のひとつは、タグである。タグは、メインコマンドの「タグ編集」機能によって設定された任意のキーワードの集合である。検索キーは、「パラメータ変更」により、随時キーワードあるいはタグいずれかのレベルに設定できる。検索レベルがタグに設定されているときには、与えられた検索キーに一致するタグがなければ、自動的にキーワードレベルで一致するキーを探索する。

さらに、検索キーのレベルに関係なく、検索キーとキーワードあるいはタグとの一致形式を設定できる。一致形式は、部分一致と厳密一致が可能である。

部分一致の場合、与えられた検索キーを含むすべてのキーワードあるいはタグと一致すると解釈される。例えば、検索キーが「平和」であれば、「平和」は言うまでもなく、「平和主義」、「平和主義者」、「平和論」、「恒久平和」、「国連平和維持軍」等々「平和」を含むすべてのキーワードあるいはタグが与えられたキーと一致すると見なされる。この形式は、検索漏れはなくなるが、検索が遅くなり、所謂ノイズが多くなる。

これに対して、厳密形式では、ワイルドカードを用いた次の4とおりの形で検索キーを与えることができる。ここでZZZZは任意の文字列、*はワイルドカード文字（利用者による設定、変更可能）とする。

ZZZZ	完全一致
ZZZZ*	前方一致（完全一致を含む）
*ZZZZ	後方一致（完全一致を含む）
ZZZZ	中間一致（他の一致形式をすべて含む）

上述の部分一致形式の指定は、この厳密形式で、中間一致を指定することに他ならないが、ワイルドカード文字を必要としない。

また、上述の検索キーのレベル、一致形式の如何に関わりなく、一度に複数の検索キーを与えることもできる。このときには、与えられた個々のキーの少なくともひとつに一致するキーをもつ文献の集合が作成される。この場合の結果は、個々のキーを指定して個別に検索を行ない、得られたすべての文献集合の論理和をとることと同じである。

3. 3 文献の表示

検索された文献、一般には任意の文献集合に属する文献は、随時画面に表示することができる。(文献集合に属する文献全体の一括ディスク出力、印刷に関しては次項を参照)。文献表示の際には、以下の機能が使用できる。

次の文献の表示

直前の文献の表示

最初の文献の表示

最後の文献の表示

文献集合中の任意の文献の文献番号、マーク番号指定による表示

表示中の文献のマーク

表示中の文献のディスク出力

表示中の文献の印刷

表示中の文献のコメントの表示、出力、入力、編集

ディスク出力、印刷は、画面表示中の文献1件について、既に設定されている出力書式に従って行なわれる。この書式は、「パラメータ変更」により、随時変更できる。出力のパラメータには、出力不要フィールドの指定、文献番号の出力如何、各フィールドごとの改行如何、文献ごとの改行如何、印刷時の行サイズ、頁サイズ、トップマージン、半角英字のフォントなどがある。

また、ディスク出力では、特に指定されない限り、PRINSEの仮定する標準出力ファイルに対して行なわれる。標準出力ファイルの名称は、必要に応じ変更できる。このファイルへの出力は、既に出力された文献などのデータを保護するため、常に追加書モードで行なわれる。

3. 4 その他の検索用機能

検索のためには、検索、文献表示の他に、以下の機能が使用できる。

これまでに作成された文献集合とそれに含まれる文献数、作成のために使用された機能等の一覧。

既成の文献集合に関してその論理積あるいは論理和となる新たな文献集合を作成する。

既成の文献集合に属する文献を著者名順に並べかえた文献集合を作成する。

特定の文献集合に関する情報をディスクに出力する。この出力は、検索時に再度入力すれば、新たな文献集合を定義することができる。

任意の文献集合を構成する文献を一括してディスク出力、あるいは印刷できる。出力先、出力書式に関しては、3. 3「文献の表示」を参照。

キーワードとタグの一覧表示（外部出力はメインコマンドにより可能）。

3. 5 コマンドモードによる検索

検索を行なうときには、上記の機能をメニューから選択して実行するだけでなく、直接コマンドとして入力し実行することもできる。また、1. 6「自動実行」で例示したように、自動実行ファイルを作成して実行するときにも、コマンドモードで実行される。コマンドモードは、メニューモードに比べ、利用者の選択の幅、自由度が高い。

検索をコマンドモードで実行するときには、上記のメニューに対応する検索用コマンドとその実行に必要な、検索キーなどのオペランドを検索用コマンドラインとして一括して入力する。このモードでは、利用者の負担を軽減するため、検索

コマンドの一覧表示、検索キーや処理対象とする文献集合の番号など必要なオペランドが欠けているときに、キーボードからの入力を促すように環境を設定する機能、予め決定できないオペランドに対して任意の文字で代用しておいて、コマンド実行時にキーボードから入力する機能、複数コマンドの一括入力（コマンドストリーム）、検索用マクロの定義、参照等の機能が用意されている。また、検索用コマンドライン（あるいはマクロ実行ファイル）に与える検索用コマンド、オペランド等では、アスキー文字の全角半角、大文字小文字の区別をしない。

コマンドストリームは、複数の検索用コマンドを指定された区切り記号で区切って与えるものであり、コマンドは与えられた順にひとつづつ実行される。次の例は、「平和」を検索して、その結果を表示するというごく単純なコマンドストリームである。前述のように入力は全角、半角を問わないし、また大文字小文字を問わない。また、この例では、Fは検索キーによる検索、Dは文献表示のコマンドであり、セミコロンがコマンドの区切り記号である。

```
F   平和   ; D   1
```

検索マクロは、頻繁に使用する検索用コマンド、文字列、コマンドストリームを登録しておき、（通常はごく短い）検索マクロを与えることにより、登録されたコマンド、文字列、コマンドストリーム等と呼出して使用するためのものである。検索マクロの入れ子も可能である。

4 データベース作成

4.1 文献データの入力

PRINSEによる検索を行なうためには、文献データをPRINSEが処理できる形に変換する必要がある。この変換処理をPRINSEへの入力と言う。新たにデータベースを作成することも、既に作成されたデータベースに追加入力することも可能である。入力媒体としては、キーボード、ディスクファイルいずれも可能で

ある。いずれの場合も、入力は何つかの手順に分割して行なう。データベースを新たに作成するときには、メインメニューの「新規作成」により、データベースの名称などを定義し、データベース用ファイルの作成、データベース情報の初期化などを行なう。以後の処理は、追加入力とまったく同じである。また、入力は随時打ち切ることができる。

4. 2 入力のパラメータ

入力に際しては、入力処理のための幾つかの記号、パラメータ確認しておく必要がある。入力処理の記号としては、複数の著者名を分離するための記号、入力データ中の注釈部分の接頭記号、置換用文字列の接頭記号、コメント行の接頭記号などがあり、環境に設定された標準値を変更することができる。入力のパラメータとしては、文献の表題からのキーワードの設定モード（ストップワードでない限りすべてキーワードとするか、既にキーワードとして登録されていない限りキーワードとしては扱わないか）、ディスクファイルからの入力時にファイル中の未処理データを別ファイルに出力するか否か、入力されたデータを別ファイルに出力するか否か、全角平仮名を単語の区切り記号とするか否か、数字で始まる語をキーワードとして扱うか否か、ディスク入力の場合、後述の編集ステップを実行するかどうかなどを確認あるいは変更する。

4. 3 入力データの編集

入力は、ふたつのステップに分けて行なう。第一は、入力データの編集であり、キーボードから文献データを入力するときには、この入力作業も含む。ディスクファイルから入力するときには、このステップは省力できるが、次のキーワード処理ステップから、このステップに戻ることもできる。

このステップは、入力されたデータの確認、編集を行なう。その他に、置換用文字列の登録、マニュアルの参照、子プロセスの実行、パラメータの変更なども可能である。

4. 4 キーワード処理

第二のステップは、キーワード処理ステップである。ここでは、処理中の文献にどのようなキーワードを付けるかを決定する。PRINSE では、個々の文献に付けるキーワードは3種類ある。著者名はすべて自動的にキーワードとする。文献の表題からは、利用者の指示に従って、抽出する。その他利用者が指定した語をキーワードとする。

表題中のキーワード設定の便を図るため、表題中の単語に関して、キーワード、異形、ストップワード、不要語の4つの状態カテゴリーを設け、それぞれ異なった色で画面表示する。異形は、複数形など表題中に出現したのとは異なった形でキーワードとして扱う語であり、ストップワードは、常にキーワードから除外することを指示した語である。不要語は、表題中に出現するが、無視する語である。各語の状態は、自由に変更できる。実際、このステップの目的は、プログラムが自動的に設定した各語の状態を変更、確定することと言ってもよい。(この方式は、基本的には、旧版の方法を踏襲している。松尾1986, 185-189参照。)

欧文の場合とはもかく、日本語文献の場合入力データを単語に分割することは不可能に近い。特に、PRINSE は既に登録されたキーワード、ストップワード、異形以外の辞書を備えていないので、単語へ分割処理はきわめて不完全である。この不備を補うために、このステップでは、単語の切り直し処理の機能を用意している。このときには、単語の先頭と末尾を指定する。

5 補助的機能

これまで述べなかった補助的な機能としては以下のものがある。最初の「索引リスト出力」と「データ編集」を除き、他の機能は、PRINSE 実行中のほぼどの局面でも随時使用できる。

索引／リスト出力

データの編集

オンラインマニュアル
子プロセス実行
環境（パラメータ）の変更，再設定

このうち、「環境（パラメータ）の変更，再設定」に関しては，1. 3に述べたのでここでは，触れない。

5. 1 索引とリストの出力

PRINSE は，本来，従来型の著者名索引，書名索引，件名索引の作成を支援するものではない。むしろこのような索引を駆逐することをこそ目的とする。しかしながら，入力した文献データの点検，修正のための機能として，以下の出力を用意している。出力先は，いずれの場合も，ディスク，プリンタが可能である。文献を出力する際には，指定された出力書式（3.3参照）が用いられる。

データベース中の全文献を出力する機能。

データベース中の全キーワードを出力する機能。該当する文献数を付けることも可能である。

データベース中の全タグを出力する機能。タグを構成するキーワードのリストを付けることも可能である。

キーワードあるいはタグを見出しとして，キーワードあるいはタグを含む文献を出力する索引作成機能。見出しとするキーワードあるいはタグに関しては，全キーワードあるいは全タグを対象とするか，画面で選択するか，ディスクからキーワードあるいはタグのリストを入力するかのいずれかが選択できる。

5. 2 データの編集と削除

データベースに収められた文献、キーワード、タグに関しては、削除が可能である。キーワードとタグについては編集も可能である。

前二者については、メインコマンドの「データ編集」で、タグに関しては「タグ編集」で行なう。

文献をデータベースから削除すると、その文献にのみ付されたキーワードも自動的に削除される。入力された文献を直接更新、修正する機能はないので、削除と再入力を組合せて行なう。

5. 3 オンラインマニュアル

PRINSE はオンラインマニュアル以外のマニュアルは提供しない。それゆえ、多くのオンラインマニュアルが、冊子体のマニュアルの要約であるのに対し、PRINSE のそれは完全なフルマニュアルである。また、この点に鑑み、オンラインマニュアルは、必要に応じ、全体あるいは部分のディスク出力、印刷が可能である。

「マニュアル」機能を選択すると、出力先の間合せがあるので、画面、ディスク、プリンタのいずれかを選択する。オンラインマニュアルは、「章」という単位から構成されており、出力先の選択に続いて、章の一覧、即ち PRINSE オンラインマニュアルの目次が表示されるので、表示、出力したい章をカーソルを移動して選択する。

マニュアルの各章は、更に「節」という単位から構成される。画面表示の際には、選択された章に含まれる節の一覧、即ち当該の章の目次が表示されるので、表示したい節を選択する。節は一般には表示画面の大きさを越えないが、カーソルにより画面をスクロールして必要な箇所を見ることができる。

5. 4 子プロセス

MS-DOS の子プロセス起動に関しては、プログラム側では特に制限はない。しかし、子プロセスの実行に必要なメモリは特に確保しないので、場合により実行できないこともある。また、多大のメモリを必要とするソフトは、640K のメモリでは、子プロセスとして実行することは困難である。

引用文献

松尾雅嗣 (1985) “パーソナルコンピュータによる平和研究文献の簡易検索システム”, 『広島平和科学』, 第 8 号, pp. 297-309

松尾雅嗣 (1986) “平和研究文献用簡易検索システム PRINSE の改良について”, 『広島平和科学』, 第 9 号, pp. 169-194

松尾雅嗣・竹花誠児 (1986) “パーソナルコンピュータによる平和科学研究センター所蔵文献の検索システム”, 『平和科学研究通信』, Vol. 9, No. 1, pp. 1-

10