

5G-4

# 検索型エキスパートシステムの 知識ベース構築支援に関する一提案

齋 礼    安信千津子    森 文彦  
( (株)日立製作所システム開発研究所 )

1. はじめに

近年、種々のエキスパートシステムが開発されつつある。これらの経験から、知識ベース構築時の問題点としては、知識の収集・整理、知識ベースのテストおよびユーザインタフェースの開発に大きな工数が掛ることが上げられている。

本稿では、検索型エキスパートシステム構築ツールESOCKS [1] を対象として、知識ベース構築作業時の問題点を解析し、より短期間で効率良く知識ベースの構築を行う方法について報告する。

2. ESOCKSの知識ベース構築作業と問題点

ESOCKSは、指定されたキーワードをもとに、文書をページ単位で検索するツールである。キーワードの大部分は検索文書から抽出される。抽出されたキーワードの整理を始めとする知識の整理には、フォーマットを固定した一連のワークシート(図2.1)が提供される。ユーザは、ワークシートに整理した知識をもとに、エディタを利用して、知識ベースファイルを作成する。

この知識ベース構築作業は、大部分が手操作、手作業であったため、記入や入力などの手作業、およびここでのミスの修正、エキスパートシステム構築ツールの文法修得にかかる時間の比率が大きい等の問題点があった。

3. 知識ベース構築支援ツール

3.1 支援ツールの構想と開発上の課題

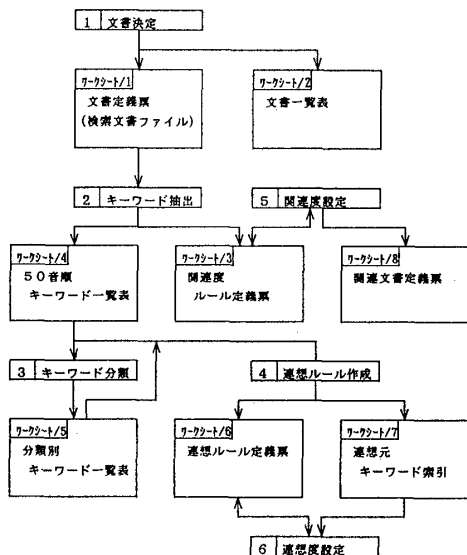


図2.1 ワークシートの関連図

ESOCKSの知識ベース構築作業は、次の3つのフェーズに大別される。

(a) フェーズ1: ワークシート作成

キーワードを抽出してワークシートに整理する作業に見られるように、人の判断と手作業とが伴う作業。

(b) フェーズ2: ワークシート間変換

入力したキーワード(連想元キーワード)から連想されるキーワード(連想先キーワード)への関係を、逆にしてワークシートに記入する場合のように、ワークシート間での知識の機械的変換作業。

(c) フェーズ3: 知識ベースファイル作成

ワークシート上の知識を、ルールやフレーム形式の知識ベースファイルに変換する場合のように、ワークシートからルールやフレームへの機械的変換作業。

以上3つのフェーズは、対象となるワークシートを媒介とするので、作業内容は独立している。そこで、各フェーズをそれぞれ効率的に行なうツール群を開発することで、知識ベース全体を効率良く構築できる。

ツール群開発の課題として、以下の点を考察した。

- ①知識ベース構築作業では、同じ文字列を何度も手操作入力することは極力避けたい。
- ②ワークシートや知識ファイルのちょっとした修正には、ワードプロセッサやエディタのような操作性の良い汎用ツールが利用できると便利である。そのために、ツールにより出力されるワークシートや知識ベースファイルを、汎用ツールの標準形式に合わせる。

その反面で、知識の内容が複数のワークシートや知識ベースファイルに関連する場合、汎用ツールを利用して1つのワークシートや知識ベースファイル上の知識だけを修正すると、残りのものに記述されている知識との間に不整合が起る。ワークシートや知識ベースファイル間の知識の不整合のチェックを、容易に行うことのできるデータ構造にする必要がある。同時に、データ構造は知識の修正(追加や削除)を容易に反映できるものでなければならない。

- ③極力エキスパートシステム構築ツールに特有の記述文法を意識しないで作業が行なえるようにする。

3.2 知識ベース構築支援ツールの概要

開発したツール群とワークシートの関係を図3.1に示す。

(1) ワークシート作成支援ツール群

検索文書からのキーワードの登録および削除、キーワード情報(50音名、分類名)の付加、キーワード分類名の登録及び削除、キーワード連想関係の設定、連想度の設定、キーワードと検索文書との関連度の設定を行なう作業を支援し、関係するワークシートへ知識の出力を行なう。

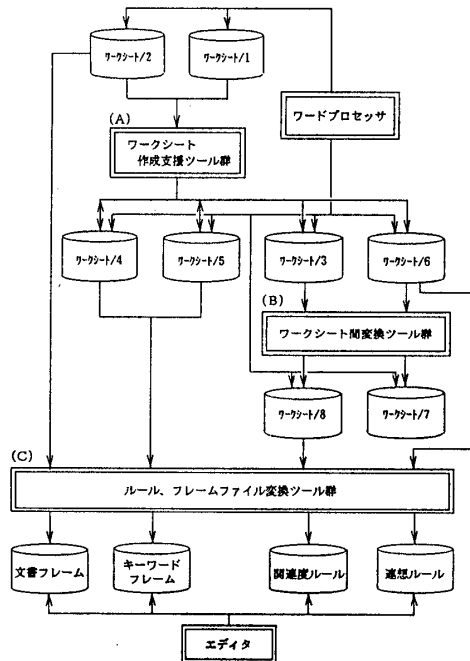


図3.1 ツールとワーシットの関係

文書タイトルなど、予めワーシットに定義した検索文書に関する知識と、(更新を行なう場合には)過去にワーシットに設定された知識(抽出キーワードなど)を読み込み、ツリー状に関連付ける。ツリーは、文書と抽出キーワードとの関係、各キーワードの属する50音やキーワード分類を表す。ツリーを探索することにより、あるキーワードが既に登録されているかどうか、どのキーワード分類に属しているか、あるいは、どの文書から抽出されたかなどの情報を容易に得ることができる。キーワードの登録や削除などの新たな設定や更新の処理は、このツリーを参照して行ない、同時に、キーワードの重複等のエラーを検出する。

設定された情報は、関係するワーシットに知識として自動出力される。出力されるワーシットは、ワードプロセッサの標準文書ファイル形式であり、ワードプロセッサを利用してその内容を確認、修正したり、プリンタに出力したりすることができる。

(2) ワーシット間変換ツール群

xからy(x,yはキーワードや文書を表す)への対応関係にある知識をもつワーシットから、逆にyからxへの対応関係を必要とするワーシットへの知識の変換、出力を行なう。

ユーザが入出力ワーシット名を指定すると、変換元ワーシットから変換先ワーシットへの自動変換を行なう。変換の際にキーワードや文書に関する登録情報を参照し、不整合を検出する。出力されるワーシットは、やはりワードプロセッサの標準文書ファイルである。

(3) ルールおよびフレームファイル変換ツール群

完成したワーシット上の知識を、各ルール、フレームの形に変換する。

知識整理の完了したワーシットから、必要な情報を取り出し、キーワードや文書に関する登録情報を参照してエ

ラーの有無をチェックし、エラーが無ければ各ルール、フレーム形式に変換して、知識ベースファイルに出力する。

出力されるファイルは、UNIXの標準ファイルであるので、汎用エディタを利用して修正を行なうことができる。また、エキスパートシステム構築支援ツールが処理できる記述文法に従って出力するので、ユーザは、ルール、フレームの記述文法について熟知する必要がなくなる。

3.3 ツールを用いた知識ベース構築手順

ツールを利用した知識ベース構築手順の概要を、図3.2に示す。ユーザは、検索文書、キーワード、50音、キーワード分類など必要最小限の情報をマウスにより指定するだけで、ワーシットやルール、フレームファイルの形で知識を得ることができ、入力やエラーチェックの面で効率的に知識ベースの構築ができる。

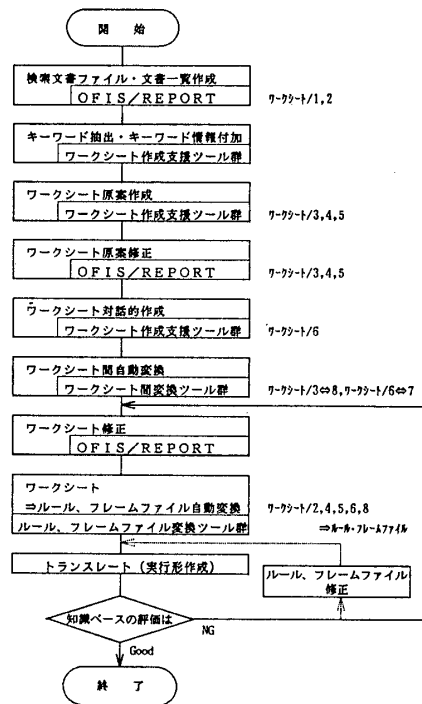


図3.2 ツールを用いた知識ベース構築手順

4. おわりに

ESOCKSの知識ベース構築を支援する3つのツール群(ワーシット作成支援ツール群、ワーシット間変換ツール群およびルール、フレーム変換ツール群)を開発した。ユーザは、これらのツールを利用することによって、ESOCKSの知識ベースを効率的に構築できる。キーワードの抽出やキーワードへの情報設定作業については数分の一、自動変換部分については数十分の一程度の時間短縮が期待できる。知識の洗練時に、繰返しツールを用いることを考えると、さらに効果が増すと予想される。

知識ベース構築支援ツールは、今後、ESOCKSのユーザに提供し、ツールの機能検証を行う。

参考文献

[1]安信、他：検索型エキスパートシステムとその基本構造に関する一考察 昭和63年度人工知能学会全国大会(第2回)、pp341-344 (1988)