

文書ベースシステム(1)

6X-1

—基本思想と表現方式—

前中 聡 田中秀俊 宮地泰造 辻 秀一
三菱電機株式会社 情報電子研究所

1. まえがき

近年、エキスパートシステム(以下ESとする)などにより一部の知識の電子化が行われ、知識の協同利用が始まった。しかし、人類が保有する多量の知識は依然文書の形態で存在しているか、もしくは個人の頭脳中にノウハウとして記憶されており、知識を共有資源として容易に活用できる環境にはなっていない。そこで、文書中に記述されている知識(以下文書知識とする)の知識ベース(以下KBとする)化を行い、知識を共有資源として容易に活用できる環境を提供する文書ベースシステム(以下DSとする)が望まれる。

2. 文書ベースシステムの概要

文書の電子化は検索の多様化、表現のマルチメディア化などを目標に研究が続けられている。我々は文書知識を記述するモデルを用意し、文書知識の蓄積ツール、利用ツールを整備して、容易に文書知識の協同利用が可能となる環境の作成を目標としている。文書に記述されている知識は様々なものがあり、文書知識を記述するモデルが複数必要となる。文書を分類し、適切なモデルを用いて文書知識を表現してシステム化することが有効である。これにより、種々の文書を自然に表現することが可能になり、文書知識の蓄積、利用が容易になる。DSはこのように様々な文書を適切なモデルによって表現された文書知識として統合して蓄積し、使いやすい文書知識の利用環境を提供することを目的とする。(図1参照)

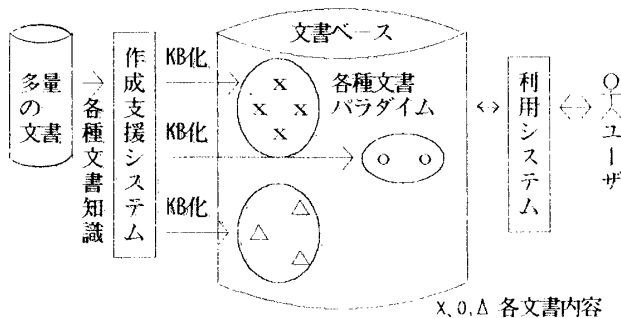


図1 文書ベースシステムの文書知識表現モデルとシステムの関係

DSは多量の文書の文書知識を統合して蓄積し、(以下、この文書知識の集合を文書ベースとする)文書知識を容易に活用可能な環境を提供する目的を実現するために、以下の2種類のツールから構成される。

- ① 文書知識を獲得、蓄積する
「文書ベース作成支援システム」
- ② 文書知識を利用する「文書ベース利用システム」
(3章参照)

上記の2種類のツールは文書知識を表現する各モデルごとに用意される。

「文書ベース作成支援システム」はモデルに従って文書知識を抽出し、文書ベースを構築するツールである。

ESは文書知識をKB化する場合、文書知識を詳細なレベル(ルール表現、フレーム表現など)まで落として表現するために作業量が多くなる。

DSの場合は文書の種類によってその内容を表現するモデルが決定されており、そのモデルに従って文書知識抽出の作業を行い、ESと比較して文書知識を文書に近いレベルで文書ベースに格納する。このため、ESと比較して文書知識抽出の作業量は少なくなる。

3. 文書ベース利用システム

文書ベース利用システムは文書ベース化された文書知識を利用するためのツールである。事例の文書知識として操作説明書の知識を選び、操作説明書のモデルを実現する文書ベース利用システム「CORPUS」をエンジニアリングワークステーション上で開発中である。

操作説明書の特徴として、

- ① 文書の内容が整理されて記述されている。
(文書知識が整理されている。)
- ② 曖昧な知識が少ない。

などがあり、最初の研究対象として適切であると考えている。文書知識を表現するモデルとして、操作対象の装置(以下装置とする)の状態遷移をネットワーク形式で表現する方式を採用している。基本機能としてネットワーク構造のノードとアークに対応している装置の状態と操作の文書知識に対する検索機能を有する。

以下に、「CORPUS」における文書知識の表現方式をのべる。

(1) 文書ベース化した文書

「CORPUS」はNC加工機の操作説明書を文書ベース化しているが、下記にNC加工機操作説明書の特徴を示す。

- ① 加工機の操作を主体に記述されている。
- ② 加工機の通常の操作をする順序に沿って記述している。(目次の構成も同じ)
- ③ 各操作に対する記述内容は具体的な操作手順と、その結果状態(ディスプレイ、操作盤の図など)である。
- ④ 各操作を実際に行う時のNC加工機の状態(各種スイッチの設定の組み合わせなど)は記述されておらず、使用者がそのデータを欲する場合はその操作の前に行う操作の操作状態結果を捜す。
- ⑤ 操作のデータをランダムにアクセスするのは困難で、ユーザは自分の操作履歴を思い出しながら操作説明書をたどってアクセスしたいデータを見つける。

(2) 操作説明書文書知識の表現手法

操作説明書の文書知識の検討を行い、「CORPUS」の文書ベースのデータ構造を決定した。

以下にその概略を示す。

- ① 操作と装置の状態(装置の操作盤上のスイッチ、ダイヤルなどの設定値、装置の可動部分の動作状態など)に着目する。
- ② 状態の変化を操作によるものと見なし、状態遷移関係をネットワーク形式(状態をノード、操作アークと見なす。)で表現している。(図2参照)

但し、装置が取り得る全操作、全状態をネットワーク形式で表現するのではなく、下記に示す有意な操作、状態のみに着目してモデル化する。

・有意な操作、状態

スイッチ、ダイヤルの設定などの詳細な操作ではなくそのシーケンスが装置の動作の指令、又はその準備などの意味のあるレベルの操作(有意な操作)に着目し、その前後の状態を有意な状態と見なしてモデル化する。

- ③ 各状態の内容は表形式で表現する。装置の状態を特徴付ける箇所(スイッチ、ダイヤル、メータ、可動部分)と、その値(スイッチ、ダイヤルの設定値、

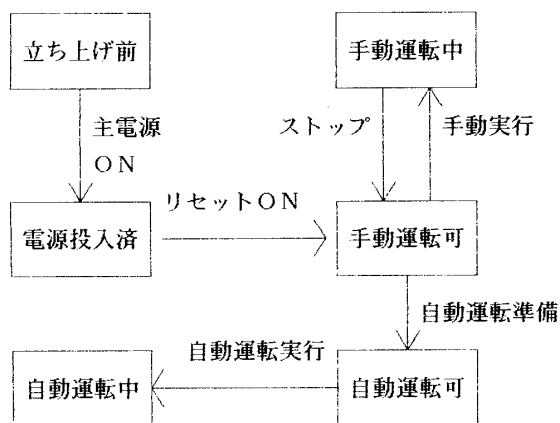


図2 NC加工機状態遷移例

メータの指す値、可動部分の動作状態など)の組み合わせを表形式で持ち、状態を決定する。

- ④ 操作の階層性(概略の操作目的から、詳細の操作手順(上記に示した有意な操作よりも詳細な操作も含む)のレベルまで)を木構造(一部ネットワーク構造)で表現している。ユーザが意識する操作目的と実際に行う操作手順の関係を表現する。

図2に示した状態遷移関係は操作、状態共に最も概略のレベルの関係を表現している。操作の階層性を状態遷移関係に反映させると、状態遷移自体が階層構造を持ち、(図3参照)装置の取り得る様々なレベルの状態と操作の関係を表現できる。これらの状態遷移関係、階層関係に対する種々の検索機能を実現することにより、様々な角度から操作説明書の文書知識へのアクセスが実現可能となる。これより、ユーザが装置の操作、状態を意識するレベルに応じた(操作目的のレベルで概略の状態遷移データを必要とする場合、操作手順そのものを必要とする場合など)文書知識の提供が可能となる。

また、CORPUSでは装置の操作ノウハウを「ある状態から他の状態への操作」、「ある操作の別手段」と位置付け、状態遷移関係のネットワーク構造に対して、ノードとアークの追加、操作の階層構造へのデータの追加を許し、ユーザのノウハウメモの文書知識の利用が可能なシステムとしている。

4. むすび

文書の形態で多量に存在する文書知識をKB化し、利用する「文書ベースシステム」について述べた。

現在、操作説明書を事例としてシステムを開発中であり、有効性を検証する予定である。

参考文献

田中他：文書ベースシステム(2) 情報第37回全国大会

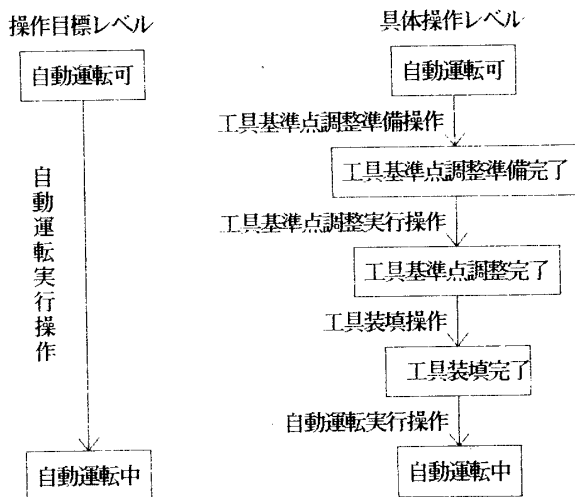


図3 NC加工機状態遷移階層構造例