

# 社会調査におけるエキスパート・システム構築について

－SSM職業コーディング支援エキスパート・システムの構想－

高橋 和子

はじめに

## I. SSM職業コーディング支援エキスパート・システムの概要

- I. 1 システムの考え方
- I. 2 システムの構成
- I. 3 タスクの型
- I. 4 開発のプロセス

## II. システムにおける知識獲得と知識表現

- II. 1 知識獲得
- II. 2 知識表現
- II. 3 実装に向けて

おわりに

## はじめに

現代社会において、職業は、人びとの基本的な社会的役割と社会的地位を決定する重要な要因である。従って、社会調査においては、性や年齢などとともに基礎的な調査項目の1つとなっている(安田・原1982)。しかし、他の属性が容易に決定できるのに比べて、職業はそれほど単純に決めることができない。なぜなら、職業は第一義的には「国民経済体系の中における労働力たる個人の、技術的活動」(安田・原1982)であるが、社会調査で仮説とされるさまざまな概念<sup>(1)</sup>に与える影響力が大きいために、その決定は慎重に行われる必要があるからである<sup>(2)</sup>。

職業は総合的に判断されるべきであるとの主張が従来よりなされてきた<sup>(3)</sup>。例えば、安田・原(1982)によれば、厳密な意味での職業を把握するには、狭義の職業(本人の仕事の内容)に産業<sup>(4)</sup>、従業先の規模(雇用者数)、従業上の地位を組合せた4次元からなる総合的な分類を考えるべきであるとする。この考えに従うと、職業は調査者自身によるアフターコーディングが必要な項目となる。なぜなら、これを決定するためには多次元からの総合的な判断が必要なため、(本人といえども)回答者や調査員に行わせることは危険だからである(原・海野1984)。

1955年から10年毎に実施されてきた「社会階層と社会移動全国調査」(以下、SSM調査(Social Stratification and Social Mobility Survey)と略称する)においては、職業は特に重要な調査項目であるために、各々3種類の自由回答(従業先の名称、従業先事業の種類、本人の仕事内容)と選択回答(従業上の地位、従業先の規模、役職名)から構成される計6種類のデータで収集された後、総合的な判断がなされる。すなわち、自由回答を含む多次元の形態で収集された職業データをSSM調査のメンバーが判断して、カテゴリーとして用意されているSSM職業分類(293小分類)<sup>(5)</sup>のいずれかにコーディングする<sup>(6)</sup>。これはSSM調査において重要かつ必須の作業であり、特にSSM職業コーディングと名付けられている。

すでに佐藤(1992)や高橋(1994b)などで指摘されたように、ここには作業量の多さと非効率性という問題が存在する<sup>(7)</sup>。従って、最近ではコーディング・マニュアル(『SSM職業分類』(原1993))をデータベース化する(佐藤1992、都築1992)、または知識ベース・システムを構築する(高橋1994b)などして、コンピュータによる支援の必要性が主張されるようになった。

前者は、いずれも回答（本人の仕事内容）に用いられた語句とコーディング・マニュアルに記述されている語句との単純なマッチングで、マニュアルを繰る代わりにコンピュータで検索しようというものである。これに対して、後者はコンピュータによる支援をさらに推し進めて、知識間の関連を利用する。すなわち、コーディング・マニュアルの記述内容や調査者の行う総合的な判断を知識としてコンピュータに持たせようとするものである。具体的には、コーディング・マニュアルの記述やSSM職業コーディングの「専門家」のもつ知識をルール型の表現形式でコンピュータに蓄え、代表的な推論機構であるプロダクション・システムを用いて、パーソナル・コンピュータ上に診断型の知識ベース・システムを構築するというものである。

現段階での状況をみると、いくつかの問題点を抱えながらも佐藤（1992）のシステムが稼働する（原1993）だけで、他のものは実用段階には至っていない。特に後者には、次のような問題が存在する。まず、システムが実際に稼働できるほどには細部の検討がなされていない。このため、専門知識の獲得方法や形式化、さらに推論機構等に関するより具体的な検討が必要である。次に、データベースを利用したシステム（佐藤1992）があるにもかかわらず、それとの関連について具体的に検討されていない。一般に知識ベース・システムを構築する場合には、既存のシステムに融合させた形態をとると有効なことがわかっている。佐藤（1992）のシステムをそのままの形で取り込まないまでも、データベースの利用という点については再考する必要があると思われる。

以上を踏まえた結果、本稿ではSSM職業コーディングにおける信頼性や効率性を高めるために、エキスパート・システムの構築を検討することを目的とする。すなわち、高橋（1994b）での結論をさらに精緻化するために、システム開発の各過程における構想内容をより具体的なレベルで明らかにする。その際、知識ベース・システムだけで閉じることをせずに、データベースと知識ベースをうまく融合させたものとしたい。

以下、次節でSSM職業コーディングにおけるエキスパート・システムの概要について述べる。3節では、この過程で特に重要な知識獲得や知識表現についての方法を具体的に述べる。最後に、4節で今後の課題を含めてまとめを行う。

# I. SSM職業コーディング支援エキスパート・システムの概要

## I. 1 SSM職業コーディングにおけるエキスパート・システムの考え方

エキスパート・システムは歴史も浅く<sup>(8)</sup>、人工知能の応用分野であるという性格も加わって、現況においては用語や内容に混乱がみられることがある<sup>(9)</sup>。本稿では、上野他(1987、1988)の定義に従い、「対象とする問題領域の専門知識を利用して推論を行い、専門的に高度な問題の解決を支援する知的問題解決システムである」と定義する。エキスパート・システムにおいては、専門家でなければ解決できないような問題を同等のレベルで解決できる能力をもつことが必要なために、専門家から専門知識を獲得することが不可欠である(上野他1987、1988)。

SSM職業コーディングを支援するエキスパート・システムは、いわゆる「職業コーディングの専門家」と同等の能力を保有して、6種類の職業データの総合的な判断により適切な職業分野コードを捜し出すシステムであると考えられる。これを可能にするには、専門家の経験や彼が参照するマニュアルの内容を知識としてコンピュータに蓄えることが必要である。

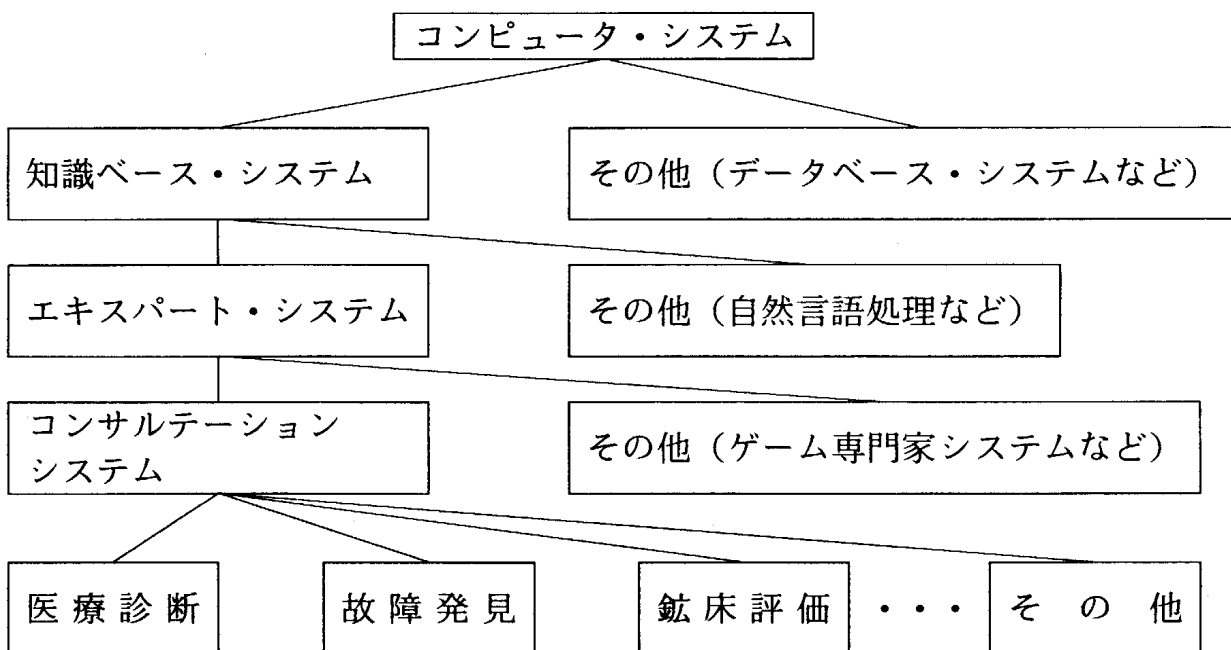
本システムを構築する目的は、主として「職業コーディングの非専門家」を支援してコーディング作業の能率を高めることであるが、それによってコーディングの信頼性や妥当性を低めるようなことがあってはならない。従って、推論の結果、複数の結論が得られた場合には一つに絞らずに、すべてを提示することとする。現在のところ、完全なシステムというものが存在しない以上、最終的な決定権はコンピュータではなく人間側に持たせることで、信頼性や妥当性の問題を解決しようとするわけである。これはちょうどワードプロセッサにおいて漢字の候補は提示するが、最後の確定は人間が行うことで、現段階での自然言語処理の不完全さを補うのと似ている。

このように、一般にエキスパート・システムは専門家に代わって非専門家の問題解決の支援を行うが、それ以外に専門家自身の問題解決も支援する場合がある。SSM職業コーディングにおいても、取扱う問題があまりにも煩雑なときには、専門家といえどもマニュアルを参照しなければ判断を誤る可能性があるわけで、エキスパート・システムの利用により作業能率や信頼性の向上を図ることができる。従っ

て、本システムは職業コーディングに関わる非専門家と専門家の両方を支援するものとなる。

ところでエキスパート・システムをソフトウェアとしてみると、知識利用型システムすなわち知識ベース・システムの種類である。Johnsonらの分類に従うと、知識ベース・システムの中でも、特に専門知識を利用するものをエキスパート・システムとよんで、自然言語処理などの一般的な知識でよいものと区別される<sup>(10)</sup>

(小橋1987)。知識利用システムに重点を置いた場合におけるエキスパート・システムの位置付けを図I.1に示す。

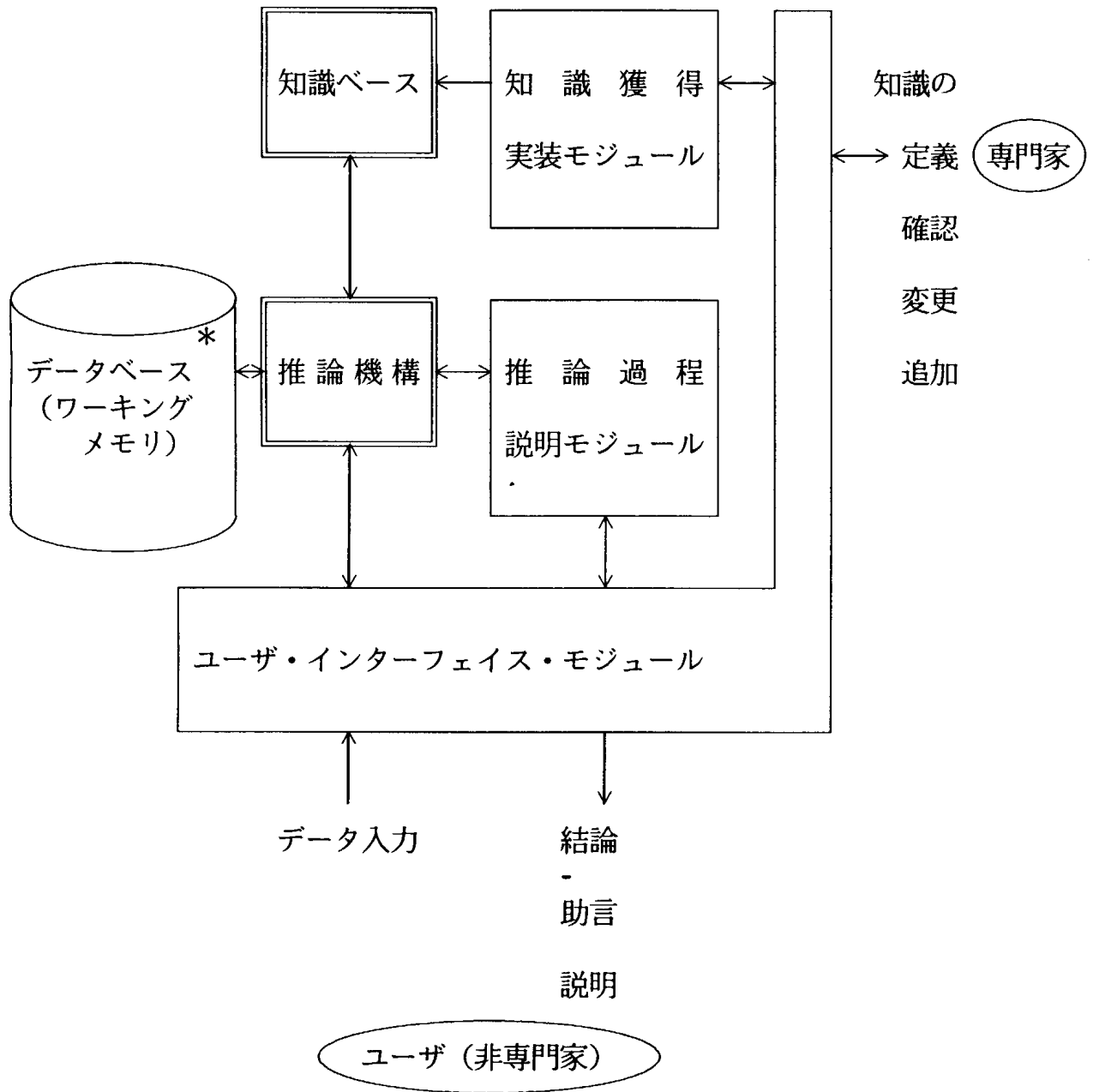


図I. 1 知識利用システムに重点を置いたエキスパート・システムの位置付け (小橋1987 : 57の一部を修正)

## I. 2 システムの構成

前述したように、エキスパート・システムは知識ベース・システムの種類であり、両者は基本的には構造が同じである。システム構成の点でエキスパート・システムを特徴づけるのは、そこで用いられる知識が特定領域の専門知識であることと、実用的なシステムを構成するという点にある。従って、エキスパート・システムは、知識ベース・システムにおける基本機能である知識ベースと推論機構に加えて、知識獲得実装モジュール (知識エディタ)、推論過程の説明モジュール、ユーザ・イ

ンターフェイス・モジュール（マン・マシン・インターフェイス）などの支援機能を持つ必要がある。まず、図 I. 2 に一般的なエキスパート・システムの基本構成を示しておこう。



\*これは知識ベース・システムにおけるデータベース（例えばワーキングメモリなど）で、一般的に用いられるデータベースとは異なる。

図 I. 2 エキスパート・システムの基本構成（上野・小山1988:8を一部修正）

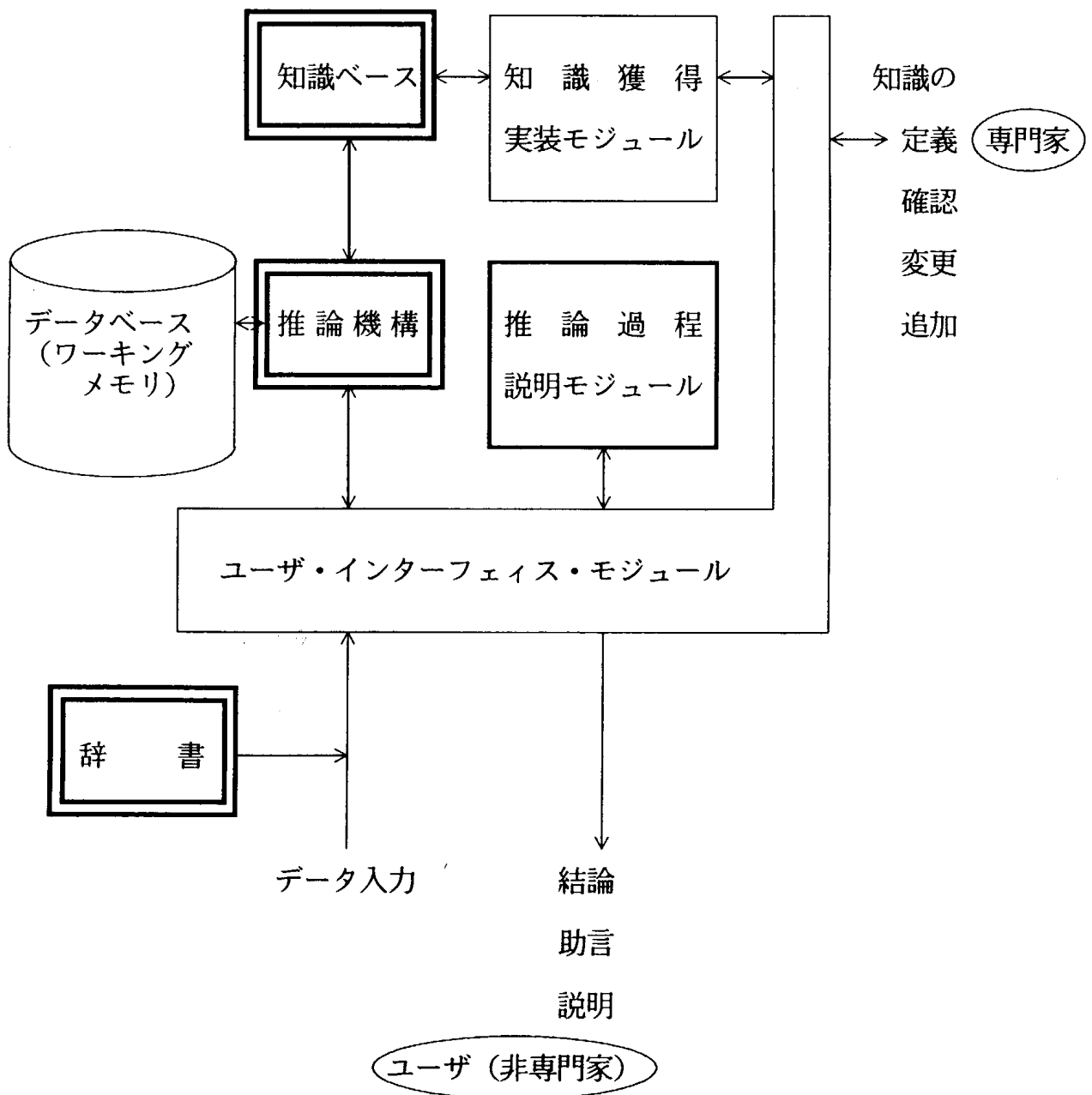
図 I. 2 における支援機能のうち、本システムにおいては推論過程の説明モジュールが特に重要である。これはシステムの透明性に関するもので、ユーザに対して推論の過程を説明する。もしエキスパート・システムが結果だけを示し、その結論を得るに至った過程を説明しなければ、ユーザはシステムの結論を納得できないであろう。SSM調査における職業の決定は特に重要であるために、コンピュータの結論を無条件に信頼することは危険であり、推論過程の説明は必要である。

ユーザ・インターフェイスは主に操作性の問題で、グラフィック表示やマルチウィンドウの利用などが考えられる。本システムは、パーソナル・コンピュータ上で稼働させるために、ワークステーション以上のものと比較すると、操作性の問題は最後まで残るものと思われる。ただし、SSM職業コーディングの性格上、ユーザはSSM調査者自身か関係する学生に限定されているために、一般的なシステムに比較するとそれほど問題はないものと判断できる。

知識エディタは、専門家から獲得した知識を知識ベースに蓄える作業を支援するモジュールである。知識ベースの容量が増したり、複雑化していく場合には重要であるが、プロトタイプの段階ではそれほど必要ないと考えられる。従って、今後の検討事項とし、今回は特に考慮しないこととする。

ところで、Winston.Pによれば、エキスパート・システムの構成原理は次の4つにまとめられる。すなわち、(1)推論機構部と知識ベース部を分離すること。(2)知識表現の形式をできるだけ均一にすること。(3)推論機構をできるだけ簡潔にすること。(4)冗長部を残してシステムを作成すること(溝口1986)。本システムにおいては、(1)から(3)については、知識表現をルール形式のみを用いることで対処できる。(4)はシステムをタフなものにするためのもので、あいまいさや不確定性に対する処理をいうが、本システムでは当面は特に考慮する必要がないものと思われる。なぜなら、SSM職業コーディングにおいては、通常、原因(回答)からみた結果(職業分類コード)は生起するかしないか(起こり得る確率が1または0である)のどちらかしかありえないと考えられないからである。ただし、対象とするものが調査データであるという制約上、原因集合を構成するものが収集されたデータでしかないという限界をもつ。従って、コードを決定する際、もし情報が不足したとしても、これ以外には何も追加できないという事態も起こり得るために、将来的には、過去のデータから算出した確率または専門家の経験などを確信度として結果に付けておくことが有効になる可能性もないわけではない。

むしろ、(4)において今回検討しておくべきことは、自然言語に関することである。本システムでは自然言語で語られる自由回答を扱う関係上、他のエキスパート・システムに比べて入力データの処理が煩雑になることが予想できる。一般的に、被調査者の回答で用いられる語句は、入力時に多少の事前編集を行うにしても、抽象度の点でコンピュータに蓄えられる語句とはかなりの隔りがあるものと思われる<sup>(11)</sup>。従って、図I. 2の構成要素では不十分で、語句の対応辞書を追加する必要がある。結局、本システムは図I. 3のような構成をとる必要がある。



図I. 3 本システムの構成 (太枠で囲まれたものは特に重要)



### I. 3 タスクの型

ここでは、本システムがエキスパート・システムにおける問題解決方法として、タスクレベルではどの型に分類されるかについて述べる。専門家が実際に問題解決を行う場合には、無数の方法が存在するといつてよい。しかし、これをタスクのレベルで考えると、分析対象の状況や結論などから分析型、設計型、教育型などその他の3つに分類できる(田中・下井1987)。

S S M職業コーディングは、前述したように、簡単にいえば被調査者の回答から職業分類コードを決める作業である。このとき、与えられた対象の状況は回答で、用意された結論は職業分類コードであると考えられる。職業分類コードはマニュアルにもあるように293個に明確に規定されており、ここでは結論の集合は確定しているとしてよい。従って、本システムは、タスクレベルではまず分析型のエキスパート・システムとして分類できる<sup>(12)</sup>。

分析型のエキスパート・システムは、分析する対象の状況や過程によりさらにデータ解釈型、診断型、制御型<sup>(13)</sup>などに分類できる。S S M職業コーディングにおいては、対象の状況(各サンプルの回答)は分析対象の属性や事実であると考えられ、またその過程は、あらかじめ用意されている仮説やカテゴリの中からこれと照合するものを選ぶことであると解釈できる。すなわち、S S M職業コーディングは用意された候補から概念や行動を割り付ける作業であると考えられるため、診断型のエキスパート・システムに分類できる。

診断型エキスパート・システムは典型的な分析型のシステムで、エキスパート・システムとしての標準的な構成を与えるものである。タスクとしては比較的単純であり、アーキテクチャも簡単なために、これまで最も多くのシステムがこの型により開発されている<sup>(14)</sup>。

ところで、診断型においては、分析対象の属性や事実(M個)は仮説(N個)と同様に有限個の集合から成ることが前提である。すなわち、この点からみれば、診断型はM対Nのマッピングと理解できよう。しかし、本システムにおいては自然言語を含んだ形態で収集されるために、Mが有限ではおさまらないことは明らかである。これを解決するためには、まず、図I. 3に示した対応辞書を充実させることが重要で、これだけでうまくいかない場合には、他の方法も考える必要がある。

なお、診断型エキスパート・システムを実現するパラダイムには、単純な仮説形成-検証モデルと仮説が階層的な分類体系を成しているとする設定/精緻化モデル

があるが、本システムは、当面は前者のパラダイムにより推論を与えていき、システムが洗練されてきた段階で、改めて後者について検討することとする。

## I. 4 開発のプロセス

エキスパート・システムは、個別的な問題を所与のコンピュータ技術、時間、費用などの限られた資源のもとで解くという点では応用的かもしれないが、理論的な基盤がなく、標準的な方法論も確立していないという批判がある（小橋1987）。しかし、エキスパート・システムもソフトウェアの一種であることから、一般には図 I. 4 のプロセスにより開発が行われる。

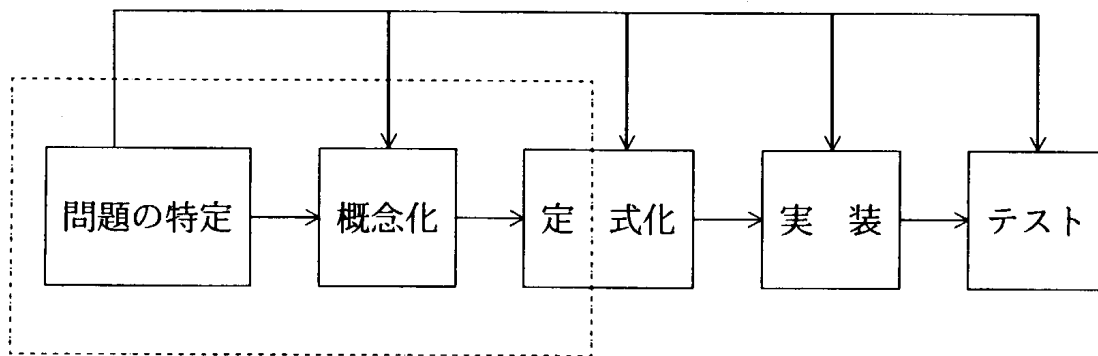


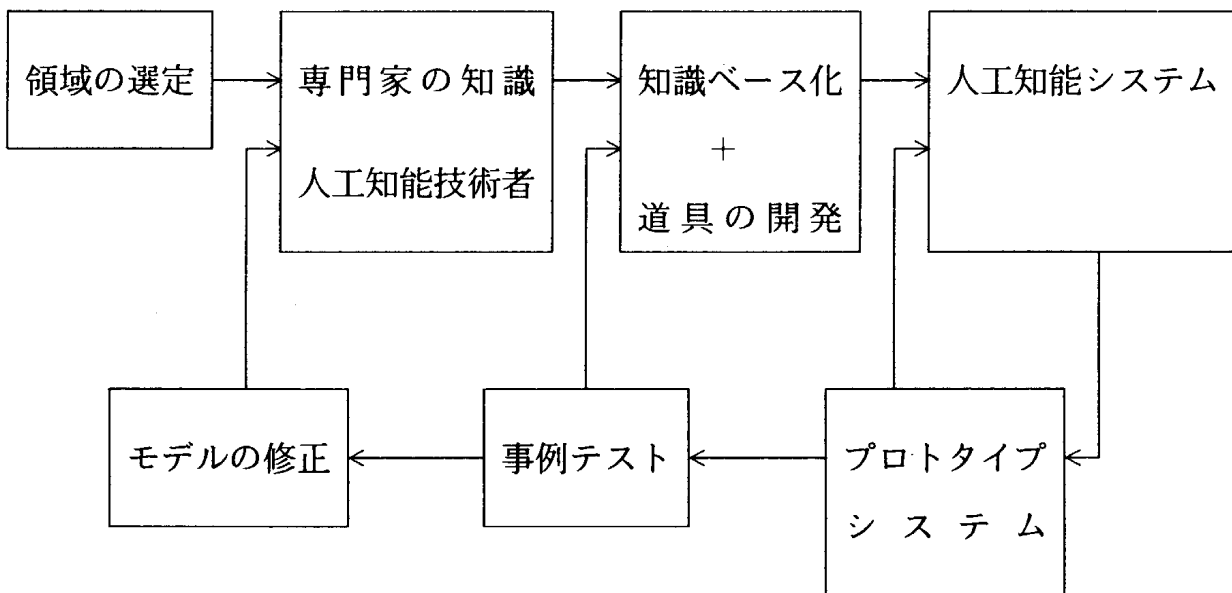
図 I. 4 エキスパート・システム開発のプロセス  
（山口（1992）を一部修正）

本稿では図 I. 4 における破線で囲まれた部分、すなわち「問題の特定」、「概念化」と「定式化」の最初の部分について述べる。各フェーズの検討に入る前に、エキスパート・システム開発における特殊性について述べておこう。

一般に、人工知能技術を用いて開発されるシステムは従来型のシステムとは異なり、螺旋型にしか構築できないことが多い。エキスパート・システムもその例外ではなく、図 I. 4 のプロセスをスムーズに辿ることが困難である。また、「概念化」は、「問題の特定」から完全に独立ではないし（折原1992）、システムの構築がかなり進んだ段階になっても、「問題の特定」に戻らざるを得ない可能性が常に存在する。これは、そもそもエキスパート・システムの開発対象が構造的に定式化できないいわゆる ill-structured であることが多いために、明確なアルゴリズムが存在しないことによる。この点で、入力と出力の関係からのみ処理方法を検討でき

る従来手法よりはるかに複雑である。

従って、実用的なエキスパート・システムを開発するには、人工知能技術によるシステムでしばしば採用される「ラピッド・プロトタイピング」を用いることが多い。これは、従来型のシステムのように、「問題の特定」を完全に解決した上で「概念化」や「定式化」に移行するのではなく、一応、動作するだけの基本知識の用意ができたらずちにシステムを作って稼働させ（実装）、あとはその挙動を調べながら「問題の特定」から「実装」までを繰り返すことで、徐々にシステムの能力を高めていくという方法である。図I. 5にそのアプローチのようすを人工知能の例で示しておく。



図I. 5 人工知能システムのアプローチ（溝口1986：11を一部修正）

ラピッド・プロトタイピングの問題点は、対象領域の専門家とシステム設計者との連絡が常に密でなければならぬために、専門家を拘束する期間が長くなることと、決定しなければならない種類の仕様を先送りにしてしまいがちなことである。しかし、従来の方法論の範囲では解決できないものも、このアプローチで開発が可能となった例もある。また、エキスパート・システムにおいては知識獲得が重要であるが、専門家は自身が保有する知識を意識しておらず、目前で実際に稼働するシステムをみて初めて必要な知識が浮き上がってくるということも多いために、メリットは大きい。

もちろんエキスパート・システムにおいても、従来型に近いアプローチがないわけではない。知識分析を重視する方法がそれで、問題解決方法の分析をできるだけ詳細に行った後に、システムの設計と実現に手をつけるものである（小橋1987）。この方法で開発できる場合には、生産性の点からもこの方法によるべきであるが、数としては少ない。

本システムにおいては、専門家との協力体制を比較的組みやすい状況であるために、できるだけ知識分析を重視しながらも、最終的にはラピッド・プロトタイピングによるアプローチの方が有効であると判断した。

さて、図 I. 4 に戻ると、「問題の特定」はシステム工学的に言えば問題定義である。これについてはすでに検討した。

「概念化」と「定式化」は、システム工学的に言えばそれぞれモデリングと設計に相当する。モデリングはソフトウェア開発の核をなすもので、現実世界の要求をソフトウェアに反映させるために、作ろうとしているシステムのモデル（仕様）を獲得し、記述する作業である。すなわち、エキスパート・システムにおいては、「概念化」は知識獲得を意味する。ただし、これは単に対象とする領域の専門家から専門知識を獲得するだけではなく、人間がもつ実世界の知識を第三者（通常は、Knowledge Engineer。以下、KEと略称する）に理解可能な言語または図表に形式化することをいう。これをコンピュータが理解できる形式に変換して（知識表現）蓄えることが「定式化」である。

変換は容易にできる場合もあるが、「概念化」における対象知識のレベルと「定式化」における知識表現言語の記述レベルの間に大きなギャップが存在する場合には困難である。従って、木下他（1987）においては対象知識を段階的に形式的知識表現に変換するために、対象概念知識構造層、知識表現モデル構造層、システム化知識層の3種類の知識階層を定義し、これを順次適用する方法を提案している（図 I. 6）。

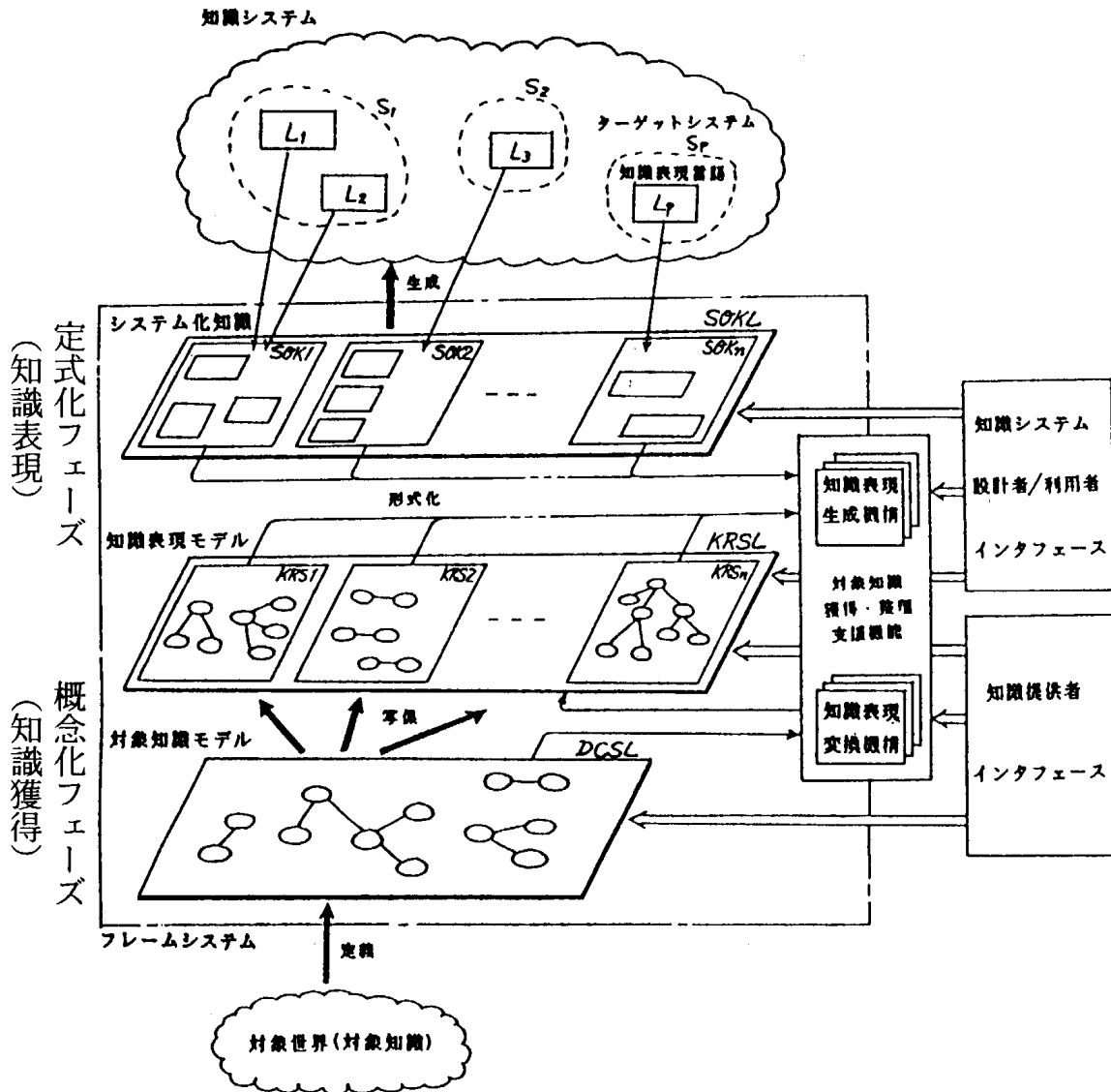


図 I. 6 多階層知識獲得支援モデルの構成  
(木下他 (1987) :144を一部修正)

厳密ではないが、本稿においては、対象概念知識構造層が「概念化」、システム化知識層が「定式化」に相当しており、知識表現モデル構造層は両者にまたがっていると解釈できる。今回は説明の都合上、「定式化」に入れることとした。

前述したように、知識獲得はエキスパート・システム開発の過程に固有なもので、最も重要なプロセスである。本稿ではこれを狭義に用いて、知識の抽出のみを意味する<sup>(15)</sup>。知識獲得は知識表現と密接に関係するために、先に知識表現形式を定めておくのが一般的である。本システムでは浅い知識しか用いないために、この表現に最も適している代表的なルール型の表現形式であるプロダクション・システムを用いることとする。

次節で、本システムにおける知識獲得と知識表現について具体的に述べる。

## II. システムにおける知識獲得と表現

### II. 1 知識獲得

エキスパート・システムにおいては、知識とは学習によって獲得され体系化されて記憶された情報であり、問題解決の過程において利用されるものである（上野他1988）。知識にはいろいろな性格のものがあり、観点の違いからいくつかに分類されるが、ここでは、浅い知識と深い知識という点に注目して考える。

本システムにおいては、対象領域において因果関係のモデルといえるようなものが存在しないために、深い知識を用いることができない。従って、浅い知識のみを用いて、浅い推論を行うこととする。一般に、診断型のエキスパート・システムは、浅い知識を用いたルール形式による推論システムとして構築できることが多い<sup>(16)</sup>。

浅い知識には、専門家の経験（ヒューリスティック）やある程度それが一般化したもの（教科書知識）があり、その獲得は通常、KEによる専門家へのインタビューや彼の問題解決過程の観察という形態で行われる。また、専門書やマニュアルなどから収集することもある。本システムにおいては、ヒューリスティックを職業コーディングにおける専門家（1人）から、教科書知識を安田・原（1982）及び『SSM職業分類』における記述からそれぞれ抽出することとした。

知識獲得はより厳密には、外部世界の知識をコンピュータにより自動的に獲得することである。本システムにおいては、検討の結果、ヒューリスティックはもちろんのこと教科書知識も自動的に獲得するのは無理であると判断されたために<sup>(17)</sup>、今回は人手で行うこととした<sup>(18)</sup>。将来的には、『SSM職業分類』については、格文法による自然言語処理技術を駆使して一部分だけでも自動的に獲得したい。

ところで、知識獲得においては、知識エンジニアリングのパラドクス<sup>(19)</sup>または知識のボトルネックなどという言葉が存在するようになりかなり困難な面がある。たとえ専門家が問題解決のために豊富な知識を持っていても、形通りのインタビューでは、エキスパート・システムを構築するに足りるだけの知識を得ることができないことが多く、方策を考える必要がある。一方でエキスパート・システムの性能は、高級

な推論機構より知識ベース・システムの充実度に大きく依存することがわかっており、この過程が最も重要となる。

本システムにおいては、専門家に対してこれまでに計2回のインタビューを行った。その概要は次の通りである（表Ⅱ. 1）。以下で詳細について延べる。

回	月日	インタビュー内容	開発フェーズ
1回目	3月22日	S S M職業コーディングの概要を尋ねる	問題の特定
2回目	5月2日	実際のデータ（100 サンプル*）に即して結果に至った過程を尋ねる	概念化

\*原・海野（1984）に掲載されている資料を用いた（資料参照）。

表Ⅱ. 1 インタビューの概要

「問題の特定」においては業務分析が中心となる。S S M職業コーディングは対象領域としては狭い範囲である上に、筆者はS S M調査メンバーの近辺にいる関係上、ある程度の知識を持っているが、担当経験がないために、必要な作業である<sup>(20)</sup>。先に図Ⅰ. 4において、「概念化」が知識獲得であるとしたが、実際には両者がそれほど明確に対応しているわけではなく、ここでも知識獲得が行われる。

「概念化」における知識獲得は、専門家の作業手順または思考手順をモデル化することが中心となる。本システムにおいては、専門家がコーディング時に適用する規則をルール化することであるが、一般に専門家の思考モデルは明確なアルゴリズムとして意識されていないことが多い（そもそも対象領域にアルゴリズムが存在しないことも多い）。従って、専門家自らが思考モデルを示すことを期待できないために、KEが実際例を提示して、それに対する彼の決定過程を聞き出すなどして、結果をまとめていく方法を探らざるを得ない。エキスパート・システムにおいて、ラピッド・プロトタイピングの方が従来型より有効なアプローチであるのは、知識獲得におけるこのような状況も大きな原因である。

本システムにおいて専門家の知識モデルを作成するためには、一般的には次のような方法が考えられよう。すなわち、職業分類コードの集計結果から多いものに的を絞り、そこに落ちてきたデータについて特徴を抽出する。それにより知識モデル

を作成すれば、少なくとも今後、その職業分類に該当する回答についてはうまく処理できるはずである。残りについても多いものを優先的にして同様な作業を繰り返すことで、徐々にシステムの機能を高めていくことができる。しかし、この方法は前提条件として、基になる調査データが調査票レベル（最低限、職業データに関して）で存在することが必要なために、不可能であることが判明した。

ところで、前出したように専門家は具体的なデータを手にした方が必要な知識が出やすい傾向がある。むしろ、彼に実際の調査データ（現存するものに限られるが）を提示して、それを職業分類コードに対応させていく過程を聞き出す方が有効ではないだろうか。結局、本システムにおいてはこの方法を採用することとした。

データとして用意したものは、原・海野(1984)に例題として掲載されている100個のサンプル（資料参照）である。ただし、本書は社会調査の講義用テキストとしての性格上、出現頻度には無関係に、典型的なものまたは誤りやすいものを中心に挙げてある。また、同じ理由により、データはSSM調査のものであるが、職業分類コードはSSM職業分類ではなく日本標準職業分類（注5参照）を用いていること、それも大分類（11分類）であることから、そのままでは本システムにおける最終的な知識モデルとすることはできないという問題がある。従って、最終的にはSSM職業小分類を目標とするものの、今回は日本標準職業大分類（以下、職業大分類と略す）のレベルで行うこととする。

専門家の思考モデルは、次に示す手順により獲得した。

①専門家が100サンプルのデータをそれぞれ職業分類（日本標準職業大分類）に対応づける過程を聞き出す（分類結果を参考資料に右に示す）。あらかじめこちらで分類した結果と異なるものについては、その理由を尋ねて明らかにする。

その結果、予想通り、単に「本人の仕事内容」だけで判断しているわけではなく、これに加えて、特に「役職名」や「従業先の規模」が頻繁に用いられることがわかった。例えば、「本人の仕事内容」が計画立案や企画、統括などとなっても、「役職名」が課長以上ではない<sup>(21)</sup>かまたはなしの場合、または「役職名」が課長以上でも、「従業先の規模」が2以下（5人未満）の場合には管理的職業従業者ではない（サンプル番号41、47、51、74、96）。

特殊な例であるが、「従業先名称」が問題になることもある。例えば水上警察であれば、「本人の仕事内容」が英語通訳であっても専門的・技術的職業従事者ではなく、保安職業従事者である（同46）。または、「役職名」が店主（販売従事者に



分類される)であっても、「本人の仕事内容」がコックであればサービス従事者である(同22)等々。

これらは安田・原(1982)や『SSM職業分類』に明示されている知識もあるが、従来より職業分類においては生産工程に従事するものを重視するといういわゆる「生産第一主義」の考え方が採られてきたという経緯による面もある。いずれにしても、単純に「本人の仕事内容」を基にデータベースによる情報検索を行うだけでは、誤った結論を導いたり、うまく探し出せない可能性があることが再確認された。  
 ②職業大分類ごとに特徴を調べるために、サンプルをまとめる(表Ⅱ. 2)。

職業大分類	サンプル数	サンプル
A 専門的・技術的職業従事者	11	23 39 42 54 56 60 66 68 79 81 91
B 管理的職業従事者	10	5 8 13 19 31 37 48 77 83 90
C 事務従事者	12	2 6 10 26 33 41 47 51 57 62 74 96
D 販売従事者	15	1 9 20 24 38 52 59 61 69 73 75 85 95 98 100
E 農林従事者	5	3 32 35 43 82
F 採掘作業	1	27
G 運輸・通信従事者	5	15 25 49 67 71
H 技能工、生産工程作業員及び労務作業員	32	7 11 12 16 17 18 21 30 34 36 40 44 45 50 53 58 63 64 65 72 76 78 80 84 86 88 89 92 93 94 97 99
I 保安職業従事者	2	4 46
J サービス従事者	7	14 22 28 29 55 70 87

表Ⅱ. 2 職業大分類別サンプル

## ③各大分類ごとに専門家の思考過程における特徴を抽出する。

ここでは、SSM職業大分類とほぼ対応する日本標準職業分類のA（専門的・技術的職業従事者）からE（農林従事者）に絞って、一般的なものと特殊なものに分けて抽出する。なお、職業を総合的に判断する際、基本的には「本人の仕事内容」を中心になされるが、この回答のされ方は一律ではなく、職種名、仕事の内容、職場名の3種類があった。

まず、一般的なものについて述べる。

## ・A 専門的・技術的職業従事者

「本人の仕事内容」の用語のみで決定されることが多い。すなわち、仕事の内容としては「研究、開発、編集、指導」であり、職種名としては「プログラマー、教員、医師」という用語が用いられる。従って、データベースによる検索で比較的うまくマッチングできそうである。

## ・B 管理的職業従事者

「本人の仕事内容」は「管理、企画、経営」であるが、前述したように、これだけで判断するべきではない。必ず、「従業先の規模」が「3以上（3～8）」であり、かつ「役職名」が「課長以上（課長、次長、部長、社長、支店長、工場長など）」であることを確認する必要がある。ただし、逆に「従業先の規模」と「役職名」が条件に合っているにもかかわらず、「本人の仕事内容」が「研究」などになっているときは、専門的・技術的職業従事者とする。これは、前述したように生産第一主義の立場による。

## ・C 事務従事者

「本人の仕事内容」は、「経理、事務、計画立案、企画、統括、設計管理」のように仕事の内容と、「庶務部人事課、経理部主計課、電子計算機係」のように職場名の2種類の回答があり、内容的にも多岐にわたる。一見すると、管理的職業従事者と思えるものも多く現れるために、これだけで判断せずに必ず「役職名」や「従業先の規模」を確認する必要がある。課長以上でなければ事務従事者である。従って、事務従事者の場合は、「本人の仕事内容」を基にしたデータベースによる検索だけでは誤る可能性が高い。なお、「事業先の規模」も確認したが、今回はすべて「3以上」であった。

## ・D 販売従事者

「本人の仕事内容」は、職種名、仕事の内容、職場名の3種類すべてがある。仕

事の内容は複数個で回答される場合が多いのが特徴である。例としては、順にそれぞれ「販売員、店員」、「事務・販売、仕入れ・小売・配達、眼鏡の調整、販売、自営、行商、集金・外交、経営全体」、「ベビー服売場」。ここでも管理的職業従事者と思われる内容があるので、「役職名」と「従業先の規模」の確認が必要である。事務従事者の場合と同様に、データベースによる検索だけでは誤りやすい。

なお、本システムでは回答が複数の職業大分類に該当する場合には全部を表示させる方針であるが、優先させるものが明らかな場合にはそれに従って選択を行う。例えば、事務従事者と販売従事者においては、後者を優先させる。

#### ・E 農林従事者

「本人の仕事内容」だけで判断する。ここでは、「いぐさ、たばこ・養蚕・水稻」のように生産品であったり、「農業」のように職種名であったりする。これは、データベースで検索できそうである。

次に、特殊なものについて述べる。ここでは、次の2例を挙げておく。

・「従業先名称」が「眼鏡店」の場合。「本人の仕事内容」が「眼鏡の調整」であっても販売従事者とする。

・同「警察署」の場合。「本人の仕事内容」が何であっても保安職業従事者とする。

以上が専門家の思考過程から得られた知識の一部であるが、これをルール形式で表現すると次のようになる。2つほど例を示す。

#### 一般的なルールの例（管理的職業従事者の場合）

もし、「本人の仕事内容」が「管理、企画、経営」であり、かつ「従事先の規模」が「3以上（3～8）」であり、かつ「役職名」が「課長以上（課長、次長、部長、社長、支店長、工場長など）」ならば、職業大分類は管理的職業従事者である。

#### 特殊なルールの例（保安職業従事者の場合）

もし、「従業先名称」が「警察署」を含むならば、職業大分類は保安職業従事者である。

ただし、これらのルールはこのままではコンピュータには理解できないために、次で知識の変換を行う必要がある。

## II. 2 知識表現

知識はいろいろな形態で獲得されるが、コンピュータで処理するには一定の形式で表現することが必要である。従って、実際にはあらかじめ表現形式を決めておい

てから知識獲得を行う方が効率がよい。本システムにおいては、前述したようにルール形式によることにしたが、これは浅い知識を表現するのに最も適している。

知識表現モデルは、(1)一様な表現(2)わかりやすい表現の2つの要件を備えていることが必要で(上野他1987)、ルール・モデル以外にも論理モデル、フレーム・モデル、意味ネットワーク・モデルなどがあり、さらにこの中から複数のを組み合わせることもある<sup>(22)</sup>。この中で、ルールによる知識表現は「IF-THEN」ルールの集合で表されるために、ヒューリスティックなど人間の思考過程をうまく表現できる。本システムにおいては、将来的にはフレーム・モデルを組込む可能性もあるが、当面はプロダクション・システムにより構築するものとする。

II. 1で作成した専門家の思考モデルを、プロダクション・ルールで表現すると、次のようになる。

一般的なルールの例(管理的職業従事者の場合)

Rule - 001

IF「本人の仕事内容」が「管理、企画、経営」かつ「従事先の規模」が「3以上(3~8)」かつ「役職名」が「課長以上(課長、次長、部長、社長、支店長、工場長など)」

THEN 職業大分類は管理的職業従事者。

Rule - 101

特殊なルールの例(保安職業従事者の場合)

IF「従業先名称」が「警察署」を含む

THEN 職業大分類は保安職業従事者

プロダクション・システムにおいては、このようなルールを知識ベースとしてもつ。例では専門家の思考モデルのみを示したが、安田・原(1982)や職業分類マニュアルにある教科書知識もルールとすることができる。前述したように、エキスパート・システムの性能は知識ベースの充実度に左右されることが多いため、実用的なシステムとするためにはできるだけ良質のルールを蓄えることが必要である。

プロダクション・システムを実現するために、本システムにおいては、OPS83を用いる。これは代表的なプロダクション・システム用の知識表現言語で、ドメイン・シェルを持つために、LISPなどのプログラム言語で開発するのに比較して、汎用性には乏しいが、目的によっては開発が容易である。

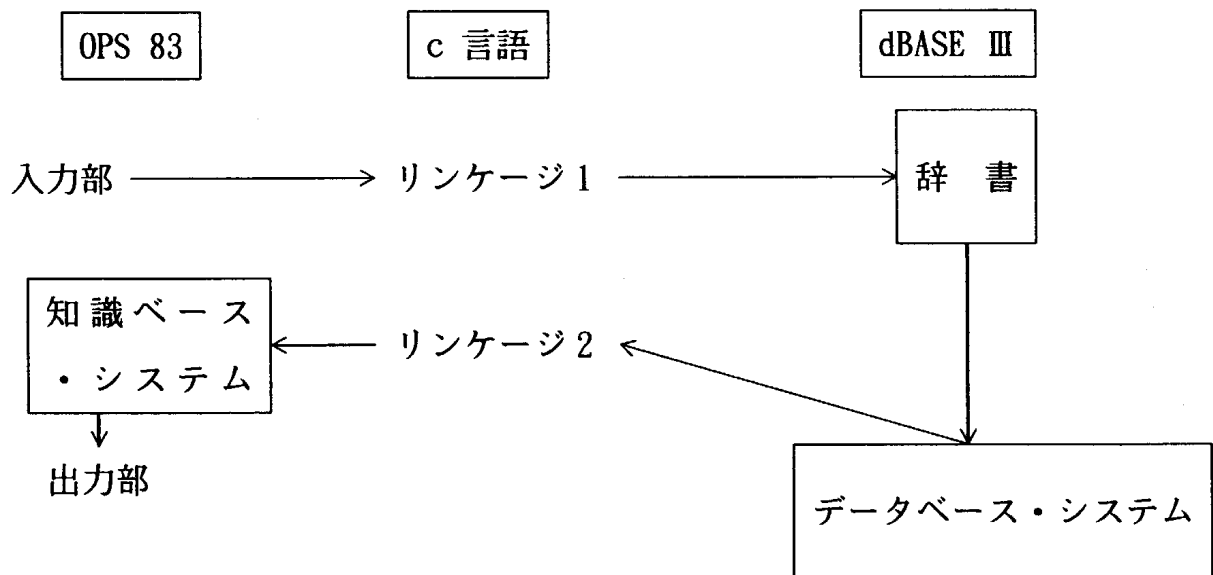
ところで、専門的・技術的職業従事者や農林従事者の場合は、図I. 3に示すよ

うな辞書の存在を条件付ければ、単なる用語のマッチングによる分類も可能になる。従って、この場合は、特にルールを作成しなくても、『職業分類』に記述されている職種名などをデータベース化して検索すればよい。また他の職業大分類の場合も、回答が職種名になっているときは、第一段階としてデータベースによる検索が使えるようである。

以上より、本システムは、システム全体としてデータベースと知識ベースを併用する形で構成するのが有効であると判断できる。最後にこの点について検討しよう。

### II. 3 実装に向けて

これまでの検討結果より、本システムは実装に向けて最終的には次のような構成をとることとする。すなわち、知識ベース部分は OPS 83、データベース部分は dBASE III による開発を行い、両者のリンケージはC言語で取る。OPS 83 は、FORTRAN やC言語などの既存の言語とのインターフェイスに優れている一方、代表的なデータベース構築用の簡易言語であるdBASE IIIは、C言語とのインターフェイスが取れる。従って、両者のリンケージをC言語で行えば、システムとしてうまく稼働できるはずである。本システムにおける構成を図II. 1に示す。



図II. 1 本システムの構成

## おわりに

本稿では、SSM職業コーディングを支援するためのエキスパート・システム構築について検討した。これにより、エキスパート・システム開発のプロセスにおいて、問題の特定、概念化、定式化までが明確になった。次には、これに基づいて実装し、実際にシステムを稼働させることになる。

前述したように、システムの開発はラピッド・プロトタイピングによるアプローチをとる。従って、プロトタイプによるシステムを稼働させて、これまでになかった新たな知見が得られるはずで、これをまた問題の特定までおろし、そこから概念化、定式化の過程をとることになる。結局、この作業を何度も繰り返すことでシステムの性能が上がるわけで、この結果については次の報告としたい。

ところで、職業データには自由回答が含まれるが、カテゴリーである職業分類は明確に定義されている。その意味で、本稿は、カテゴリーが存在する場合における自由回答の処理・分析方法を模索するための一考察としての位置付けもなされよう。今後は本システムの実装化を目指すと同時に、類似の分野についても検討を行いたい。

### 謝辞

知識獲得段階において、1985年SSM調査の代表的なメンバーで、職業コーディングの専門家でもある原 純輔東北大学教授に多大な御協力を頂いたことを記して感謝します。

### (注)

- (1) 職業は、生活様式、社会関係、社会的態度、パーソナリティなどに、大きな影響を与えらるゝと考えられる(原・海野1984)。
- (2) 官庁統計や一般の調査においては、あらかじめ選択肢で職種名や仕事内容などを提示して回答させており、比較的簡単に決めている。目的によってはこれで十分なこともあるが、社会調査においては適当な方法であるとはいえない。
- (3) 例えば、狭義の職業に従業上の地位だけを組合せたもの(昭和28年住宅統計調査)や、安田三郎による狭義の職業、従業上の地位、企業規模を組合せたものなどが提唱されている。

- (4) 産業とは、ここでは原則として、事業所において業として行われる経済活動をいう。また、教育、宗教、公務、非営利団体などの諸活動も産業に含ませる（安田・原1982）。
- (5) S S M職業分類は、現在、わが国で用いられている最も詳細な職業分類である行政管理庁の「日本標準職業分類」に基づいて作成されたものである。日本標準職業分類における分類体系は次の通りである。

大分類	中分類	小分類
A 専門的・技術的職業従事者	11	64
B 管理的職業従事者	3	9
C 事務従事者	4	10
D 販売従事者	2	13
E 農林従事者	3	16
F 採掘作業	2	10
G 運輸・通信従事者	5	25
H 技能工、生産工程作業員及び労務作業員	22	189
I 保安職業従事者	1	9
J サービス従事者	3	25
K 分類不能の職業	1	1
計 11	57	371

S S M職業分類の体系は次の通りである。両者の最も大きな違いは、日本標準職業分類におけるHがS S M職業分類では5、6、7に分割されている点である（含F）。また、CとI、Jは3に統合されている。

大分類	小分類	(日本標準職業分類)
1 専門的職業	48	(A)
2 管理的職業	10	(B)
3 事務的職業	21	(C)
4 販売的職業	17	(D)
5 熟練的職業	81	(H)
6 半熟練的職業	72	(H)
7 非熟練的職業・単純労働者	26	(H)
8 農林的職業	13	(E)

9 その他（分類不能など）	6	(K)
計	8	294

- (6) 職業小分類コードはそのまま使用される場合もあるが、より大きなものにまとめられることもある。例えば、自営業者および家族従事者、専門職業従事者、管理的職業従事者、大企業ホワイトカラー労働者、中小企業ホワイトカラー労働者、大企業ブルーカラー労働者、中小企業ブルーカラー労働者、農業および類似職業従事者の8分類がある。
- (7) SSM調査におけるサンプル数は約4000で、職業データに関しては、本人の職歴、（結婚している場合）相手の職歴、父親の職歴など1サンプル当たり平均10個の職業が尋ねられるために、約40000個についてコーディングをする必要があった。コーディング作業にはSSM調査のメンバー20人～30人×7日を費やした。
- (8) エキスパート・システムの研究は、1965年にスタンフォード大学のFeigenbaumらによる化学構造式同定システムDENDRALに端を発する。70年代は、主としてアメリカの大学や研究機関で研究開発が行われた。当初は個別的応用志向であったが、後半になると汎用化された知識表現言語が開発され始め、80年代に入って産業界へも波及し始めた。我が国では1982年にICOTの設立により、大学や産業界での研究が盛んになり今日に至る。
- (9) エキスパート・システムの一環として、人工知能の基礎研究である専門家の認知モデル、深い知識構造、知識獲得、対象モデルなどを指すことがある反面、人工知能技術の簡単な応用、例えばIF-THENルールで書かれただけのプログラムをエキスパート・システムと称している場合もある。もっとも応用といっても、小規模のものから大規模なものまであるが、後者の知識ベース・システムの開発などは、人工知能研究の質的向上にとって非常に重要である（上野1988）。
- (10) この理由から、高橋（1994b）では知識ベース・システムとしたが、本稿ではこの点を強調してエキスパート・システムとした。
- (11) 例えば佐藤（1992）によると、回答では「海であわびを取っている」が、マニュアルでは「海草・貝採取人」である。
- (12) これに対して、設計型のエキスパート・システムでは結論の集合が確定していないために、要求仕様に合わせて構成要素を組合せていかなければならない（田中他1987）。
- (13) 制御型のエキスパート・システムは、準備された候補の中から適切な制御動作を決定する意味で分析型に分類できるが、推論や動作の実時間性が厳密に要求される点で、診断型と異なる（前掲書）。
- (14) エキスパート・システム概念を確立した感染症の診断システムMYCINが代表的なも



のである。

- (15) 知識獲得が広義に用いられることもある。この場合には、知識の抽出だけでなく、知識の変換、知識の構造化・洗練化を含む。
- (16) これに対して、設計型のエキスパート・システムにおいては深い知識を用いた深い推論が必要なことが多い。
- (17) 高橋 (1994a)において、職業分類マニュアルの構造化を格文法により行おうと試みた結果、自動的に知識を抽出する可能性は認められたものの、現段階では無理であると判断された。
- (18) 一般的にも、知識表現や知識の利用が人工知能技術としてある程度確立しているのに対して、知識獲得はまだ研究段階にある。
- (19) 知識のパラドクスとは、Watermanにより名付けられたもので、問題領域の専門家が優秀であればあるほど、自分の問題解決に使う知識を記憶することが不得手になる傾向があることを意味する。
- (20) またインタビューの前に、SSM調査の結果により刊行された行書（『現代日本の階層構造』全4巻（東大出版会））の関係箇所を目を通しておくなどして、対象領域に対する知識を得るようにした。
- (21) 係長、主任、主事などは課長以上には該当しない。
- (22) 論理モデルは大きくはルール・モデルに属するもので、1階述語論理の体系において知識を表現し、三段論法によって推論を行う。厳密な論理展開を表現できるために深い知識を扱う場合に向く。この点で、本システムで用いた浅い知識を扱うプロダクション・システムと対比的である。Prologとの相性がよい。フレーム・モデルはMinskyのフレーム理論に基づく知識表現モデルで、人間の記憶や認識の心理学モデルを統一的理論として体系化したものである。ある一定の概念対象を表現するためのデータ構造をもつ。プロダクション・システムを始めとする他のモデルと併用されて効果が上がることも多い。本システムにおいても、将来的には検討したい。意味ネットワークは大きくはフレーム・モデルに属すると考えられるもので、ネットワークに特別の意味をもたせるようにしたデータ構造である意味ネットワーク技法を用いて知識を表現するモデルである。ただし、標準的な意味ネットワークというものはなく、ノードとアークを用いてネットワーク全体が特定の意味をもつ知識体系を総称している。

[参考文献]

- C.L.フォギー.1986.『人工知能用言語OPS83』（中間正人訳）パーソナルメディア
- Creedy, H.R. et al. 1992. Trading MIPS and Memory for Knowledge Engineering. *Communications of the ACM*. Vol.35, No.8:48-63.
- 福田収一他.1990.『OPS83プログラミングテクニック』パーソナルメディア
- 原 純輔.1986.「職業移動のネットワーク」『リーディングス日本の社会学8 社会階層・社会移動』（直井優他編）:214-228 東京大学出版会
- .1993.『SSM職業分類（改定版）』（文部省科学研究費補助金研究結果報告書）
- ・海野道郎.1984.『社会調査演習』東京大学出版会
- 秦 勝範.1991.『エキスパート・システムとシステム・エンジニア』工学図書
- 飯島 正他.1994.「実世界と形式的記述の接点としてのオブジェクト指向モデル—周辺分野のラウンド・トリップを通して—」『情報処理』Vol.35, No.5:412-422
- 近畿情報処理技術士会編.1989.『システムエンジニアのためのエキスパート・システム入門』電波新聞社
- 木下哲男他.1987.「多層階層知識獲得支援モデル」『人工知能システムの枠組みシンポジウム論文集』情報処理学会
- 小橋康章.1987.「知識抽出の方法と問題点」『知識の獲得と学習』オーム社（大須賀節夫他編）:53-72
- 古在豊樹他.1987.「園芸環境工学における最近の話題(12) 農業用エキスパート（コンサルテーション）システムの開発と応用①」『農業および園芸』Vol.62, No.2:342-348
- .1987.「園芸環境工学における最近の話題(13) 農業用エキスパート（コンサルテーション）システムの開発と応用②」『農業および園芸』Vol.62, No.3:439-445
- 溝口文雄.1986.『AIテクノロジー』オーム社
- 尾方 説.1992.「知識エンジニアにおける知能獲得の実際」『情報処理』Vol.33, No.6:620-622
- 折原良平他.1992.「仕様獲得 VS.知識獲得」『情報処理』Vol.33, No.6:605-611
- Fikes, R. and T. Kenhler. 1986. The Role of Frame-Based Representation in Reasoning. *Communication of the ACM*. Vol.28, No.9:904-920.
- 佐藤嘉倫.1992.「職業コーディング支援システムの構築」『非定型データの処理・分析法に関する基礎的研究』（代表 原純輔）文部省科学研究費補助金研究成果報告書:199-204
- 高橋和子.1994a.「職業分類における自然言語処理について—知識ベース・システム構築のための予備的考察—」『千葉敬愛短期大学紀要』Vol.16:127-140

———.1994b.「職業コーディングにおける知識ベース・システム構築の構想」

『千葉敬愛短期大学環境情報研究』Vol.2:77-88

田中 博・下井優一.1987.『エキスパートシステム構築の方法』パーソナルメディア

都築一治.1992.「職業コーディング自動化システムの試験的構築」『非定型データの処理・分析法に関する基礎的研究』（代表 原純輔）文部省化学研究費補助金研究成果報告書:205-214

上野晴樹・石塚 満.1987.『知識の表現と利用』オーム社

———.小山照夫.1988.『エキスパートシステム』オーム社

山口高平.1992.「知識獲得技術はソフトウェア開発方法論に適用可能か？」『情報処理』Vo 1.33, No.6:623-630

安田三郎・原 純輔.1982.『社会調査ハンドブック』第3版. 有斐閣

資料

第3回SSM調査結果（職業データを中心にしたもの）（原・海野(1984)）

番号	年齢最終		従 業 先			職 業			職業分類
	(歳)	学歴	名 称	事 業 内 容	規模*	地 位	仕 事 内 容	役 職	コード
001	55	中学	〇〇農産カウスタ	ガソリン販売	3	雇用者	販 売 員		D
002	38	高校	〇〇精機	カメラ部品製造	4	雇用者	経 理		C
003	67	不明		農 業	1	自 営	米 作		E
004	28	高校	〇〇警察署〇〇駐在所	警 察	8	雇用者	警 察 官		I
005	41	中学	〇〇食品株	食肉加工業	3	雇用者	工場の管理全般	工 場 長	B
006	27	大学	〇〇インフィルコ株	下水処理のプラント設計	7	雇用者	総務部人事課		C
007	37	中学	〇〇製鉄	製鉄業	4	雇用者	精 選 工		H
008	45	大学	〇〇工機	工業機械の販売	5	雇用者	企 画	部 長 代 理	B
009	25	高校	〇〇青果	果物の委託販売	5	雇用者	事務・販売	主 任	D
010	52	高校	〇〇県庁		8	雇用者	一般事務	主 事	C
011	50	高校	〇〇織物	裏着製造販売	2	自 営	いっさい		H
012	34	高校	〇〇サイクル	自転車販売	1	自 営	販 売 修 理		H
013	42	大学	〇〇光学機械K K	光学精密機械製造	7	雇用者	経理部主計課	主 計 課 長	B
014	28	高校	レストラン〇〇〇〇〇	飲 食 店	3	雇用者	ウエイター		J

番号	年 齢 最 終		従 業 先			職 業			職業分類
	(歳)	学歴	名 称	事 業 内 容	規模*	地 位	仕 事 内 容	役 職	コ ー ド
015	41	中学	国鉄〇〇機関区	鉄 道 業	7	雇 用 者	貨物列車運転	機 関 士	G
016	41	中学	〇〇ポンプ	ポンプ製造	5	雇 用 者	旋 盤 工		H
017	38	中学	〇〇製作所〇〇工場	モーターのケース (鋳物)	7	雇 用 者	造形工 (鋳型)		H
018	33	高校	〇〇製作所	鉄 加 工	4	雇 用 者	機 械 加 工		H
019	43	高校	〇〇商工リサーチ	興 信 所	7	雇 用 者	庶 務	部 長	B
020	24	大学	〇〇ストアー〇〇店	スーパーストア	5	雇 用 者	ベビー服売場		D
021	39	大学	〇〇建設工業	電話線の敷設	7	雇 用 者	電話線工事	工 事 長	H
022	46	高校	〇〇食店	中華料理店	3	自 営	コ ッ ク	店 主	J
023	56	大学	〇〇〇中学校	中 学 校	8	雇 用 者	校 長	校 長	A
024	50	大学	〇〇商店	日用品・食料品の販売	1	自 営	仕入 小売 配達		D
025	43	中学	〇〇産業	砂利運搬 販売	4	雇 用 者	ダンプカーの運転		G
026	44	大学	〇〇信用金庫	金 融 業	7	雇 用 者	電子計算機系		C
027	43	中学	〇〇石炭〇〇鉱業所	石炭採掘業	7	雇 用 者	採 炭 員		F
028	30	中学	〇〇理容室	理 容 業	3	雇 用 者	理 髪 師		J
029	54	中学	〇〇病院	精神病院	4	雇 用 者	調 理 師		J
030	63	中学	〇〇研究室	建築用鉄筋の強度検査	3	雇 用 者	検 査	室 長	H
031	33	高校	〇〇電器	コンデンサ製造	6	雇 用 者	製 造	課 長	B
032	45	中学		漁 業	1	自 営	漁 師		E
033	52	高校	〇〇〇〇〇エレクト工業	生こん製造	3	雇 用 者	料 理		C
034	62	中学	〇〇洋服店	洋服仕立販売	1	自 営	裁 断	店 主	H
035	51	中学		稲 蘭 草	1	自 営	稲 蘭 草		E
036	40	中学	〇〇石材店	墓・石碑造	3	重 役	塑 刻		H
037	55	高校	〇〇営林署	国有林の管理	8	雇 用 者	庶 務	庶 務 課 長	B
038	22	大学	〇〇眼鏡店	眼鏡の販売	2	家 族	眼鏡の調整		D
039	24	高校	〇〇コンピュータシステム	データ調べ	4	雇 用 者	プログラマー		A
040	25	高校	〇〇自動車工業所	自動車の板金塗装	3	雇 用 者	塗 装		H
041	25	大学	(株) 〇 〇	広 告 業	7	雇 用 者	広告計画立案		C
042	30	大学	〇〇社東京本社	出 版 業	5	雇 用 者	編 集		A
043	43	中学		農業 (りんご, 大豆, 米)	1	家 族	農 業		E
044	61	高校	〇〇商事	サービス業	4	雇 用 者	食堂などの雑役		H
045	32	中学	〇〇金属	金属プレス	3	雇 用 者	金属プレスと外交		H
046	22	大学	〇〇水上警察署		8	雇 用 者	英語通訳		I
047	51	高校	〇〇装置	プラント工事組立	4	雇 用 者	企 画		C
048	54	大学	〇〇〇〇〇〇センター	雑役下請業	3	重 役	経 営	社 長	B
049	36	高校	〇〇水産試験部	漁業および水産物加工	7	雇 用 者	二等航海士		G
050	31	中学	〇〇工務店	大 工	2	自 営	大 工		H
051	49	高校	〇〇〇〇区役所		8	雇 用 者	公園建設事務	建 設 係 長	C
052	66	中学	〇〇商店	菓子, 雑貨販売	1	自 営	販 売		D
053	27	高校	〇〇計器	ガス器具等の販売	4	雇 用 者	サービスマン	主 任	H
054	34	大学	電 電 公 社	電報・電話	7	雇 用 者	架線工野郎	主 事	A
055	32	中学	〇〇学園高校		5	雇 用 者	調 理 師		J
056	34	大学	〇〇大学化学研究所	大 学	8	雇 用 者	抗癌物質の研発合成	助 手	A
057	56	大学	〇〇港管理事務所	港湾管理 施設貸与	8	雇 用 者	野積場の使用許可など		C
058	46	中学	米穀商工業協同組合	米の販売	3	雇 用 者	米の加工		H

番号	年齢最終		従業先			職業			職業分類
	(歳)	学歴	名称	事業内容	規模*	地位	仕事内容	役職	コード
059	58	中学	〇〇家具店	家具小売店	2	雇⽤者	店員		D
060	47	高校	〇〇市立〇〇小学校	小学校	8	雇⽤者	教職		A
061	38	高校	〇〇市農業共済組合	農業保険業務	4	雇⽤者	保険契約外勤		D
062	30	高校	〇〇縫製KK	婦人服縫製	4	雇⽤者	営業事務	係長	C
063	29	高校	〇〇製作所	ガスメーカー製造	6	雇⽤者	資材管理	倉庫班長	H
064	26	高校	〇〇工業	自動車修理業	3	雇⽤者	板金工		H
065	63	中学	〇〇製糖〇〇工場	缶詰用缶製造	6	雇⽤者	品質検査	嘱託	H
066	40	大学	〇〇高等学校	高校	8	雇⽤者	高校教員		A
067	30	中学	〇〇油化	化学工場	7	雇⽤者	運転員		G
068	47	大学	〇〇〇化学	工業薬品製造	5	重役	薬品の研究	取締役	A
069	31	中学	喫茶〇〇	喫茶店	2	自営	自営		D
070	25	高校	〇〇相互銀行	相互銀行	6	雇⽤者	受付		J
071	37	中学	〇工業	特殊機械(カクカ等)貸付	4	雇⽤者	機械の運転		G
072	23	高校	〇〇精工〇〇支社	自動車部品販売	7	雇⽤者	旋盤		H
073	64	中学		売薬行商	1	自営	売薬行商		D
074	30	大学	〇〇〇工業	自動車部品製造	7	雇⽤者	企画		C
075	39	中学	〇〇木材	木材売り	3	雇⽤者	集金 外交		D
076	50	高校	〇〇靴店	靴屋	1	自営	製造 販売		H
077	28	大学	〇〇ベビー	貸おむつ業	4	雇⽤者	営業管理	営業次長	B
078	43	中学	〇〇社	製本業	1	家族	製本工		H
079	51	大学	〇〇〇病院	精神病院	4	雇⽤者	精神医師		A
080	35	高校	〇〇製鋼	帯鋼、パイプ製造	7	雇⽤者	焼鈍工		H
081	36	高校	〇〇放送〇〇支社	放送局	6	雇⽤者	技術専門職	主事	A
082	53	中学		煙草 養蚕 水稻	1	自営	煙草 養蚕 水稻		E
083	46	大学	国民金融公庫〇〇支店	政府金融機関	7	雇⽤者	管理職	支店長	B
084	35	高校	〇〇鉄筋〇〇出張所	工事の鉄筋組立	3	雇⽤者	現場監督		H
085	56	高校	〇〇理化器械	理化学用器機械販売業	2	自営	経営全般	代表取締役	D
086	27	高校	〇〇自動車工業所	自動車の板金修理	3	雇⽤者	修理		H
087	40	大学	〇〇食堂	食堂	1	自営	調理師	経営者	J
088	27	高校	〇〇〇〇モーター	自動車販売 修理	6	雇⽤者	整備		H
089	25	高校	〇〇堂	パン、洋菓子の製造 販売	4	雇⽤者	製造		H
090	41	高校	〇〇〇ヤ	紳士婦人服縫製販売(卸、小売)	4	雇⽤者	製造部管理	製造部次長	B
091	33	大学	全国漁業組合連合会	漁業組合	6	雇⽤者	漁業用機器の指導	技術主任	A
092	32	高校	〇〇工業所	溶接業	2	雇⽤者	熔接工		H
093	30	中学	〇〇精器	機械製造	4	雇⽤者	研磨	班長	H
094	36	高校	〇〇自動車〇〇工場	乗用車の製造	7	雇⽤者	生産管理	総括(係長)	H
095	57	中学	〇〇環苑KK	公害防止機械の販売	2	自営	販売	社長	D
096	54	高校	〇〇〇市役所	市役所	8	雇⽤者	建築の設計監理	係長	C
097	62	中学	〇〇織物	帯地製造	4	雇⽤者	帯地製造		H
098	39	大学	〇〇新聞店	新聞店	2	自営	店主		D
099	46	中学	〇〇工務店	建築業	2	自営	経営主		H
100	65	高校	〇〇〇屋	袋物卸	4	重役	販売	重役	D

\* 従業先の規模は次の通りである。

1 : なし      2 : 1人以上      3 : 5人以上      4 : 30人以上      5 : 300人以上      6 : 500人以上      7 : 1000人以上      8 : 官公庁

## Construction of an Expert System for SSM Occupation Coding

Kazuko Takahashi

In social surveys, although occupation coding is very important, it is also very difficult. In this paper, I propose an expert system which supports SSM(Social Stratification and Social Mobility) occupation coding.

The system contains both a knowledge-based system and database system. One can be developed by OPS83 and the other by dBASEIII which is linked by the programming language C. It uses the production system which is a kind of rule type inference.