

平均圧縮情報量を用いた事例ベースからの知識発見

指導教員 皆川 勝

学生氏名 上谷 丈和

エキスパートシステムの開発にあたっては、知識獲得が困難であり、またあいまいさを持った情報の取り扱いが必要である。本論文では、鋼橋疲労損傷の補修方法選定を対象問題として、皆川らによるルールベース洗練機能付推論システムが対象問題によらず機能するかを検討した。さらに、データベースからの知識発見 (KDD) の考え方を用いた平均圧縮情報量によるルール抽出アルゴリズムを提案し、知識獲得問題解決のための一手法を提案した。この結果、皆川らによる推論システムは、対象問題によることなく推論システムとして十分有効に機能していることが示され、また、平均圧縮情報量がデータベースからの知識発見に有効であることが示された。

Key Words: knowledge discovery, expert-systems, knowledge acquisition, case-bases, network-systems

1. はじめに

エキスパートシステムを構築する上で、最も困難なプロセスは、専門家や熟練技術者が持つ問題解決能力や経験的知識を明文化(ルール化)する、知識の獲得作業である。これを知識獲得問題と呼んでいる。知識獲得は、システム構築の初期における専門家の知識の体系化と、知識がある程度獲得された後の、知識の追加・修正の過程に大きく分けられる^{1),2)}。

システム構築の初期においては、一般に、専門家や熟練技術者へのアンケート調査等から、対象問題の主な特徴点を抽出してルール化し、ルールベースを構築する。この段階では、手作業によりルールベースを構築することから、多大な労力を要する困難な作業となる。また、知識の追加・修正においては、あるルールが表現すべき因果関係に、ルールベース内において矛盾が生じたり整合を取れなくなる場合があり、いかにルールベース内の整合性を保ち、信頼性の高い評価をしてルールの追加・修正を行うかが問題となる。

一方、専門家や熟練技術者の示す知識の多くが、経験則に基づく経験的知識であることから、示される知識は、複雑なものであったり、曖昧さを含み真偽が明確でないものであったりする。

このように、ルールの追加・修正を含めた知識獲得問題は、現在においてもエキスパートシステムを構築する

際の最も重要な課題の一つとなっており、土木分野においても、知識獲得問題に関する研究がなされている。

宮本らは、対象問題をコンクリート橋診断としたエキスパートシステムの開発を行っている。その中で、ファジィ集合論を用いた知識の更新手法の提案³⁾や、経験的知識に内在する曖昧さをファジィエントロピーの概念を適用して定量的に評価するアルゴリズムを提案⁴⁾している。

三上らは、対象問題を鋼橋疲労の補修方法選定としたエキスパートシステムの構築を行っている。ここでは、推論エンジンにニューラルネットワークの学習機能を持たせ、知識の追加、修正及び削除を可能にし、さらに類推推論機能と負の学習を考慮する機能も実装した^{5),6)}。また、三上らは、事例ベース推論を用いた、鋼道路橋疲労損傷の補修方法選定も報告している^{7),8),9)}。

また、皆川らは仮説をノードとし、仮説間の関係を表す結合で構成される相互結合型のネットワークでルールベースを構成し、知識洗練機能付き推論システムを構築して、これを道路橋 RC 床版の損傷要因推定問題に適用してその有効性を示した¹⁰⁾。

一方、知識獲得の初期における知識のルール化に対して、決定木及び遺伝的アルゴリズムによる知識獲得手法¹¹⁾、帰納学習を用いた帰納推論によるルールの獲得手法¹²⁾や、事例ベース推論を用いた知識獲得手法¹³⁾⁻¹⁵⁾などによる、知識獲得問題の解決のためのアプローチが、知識

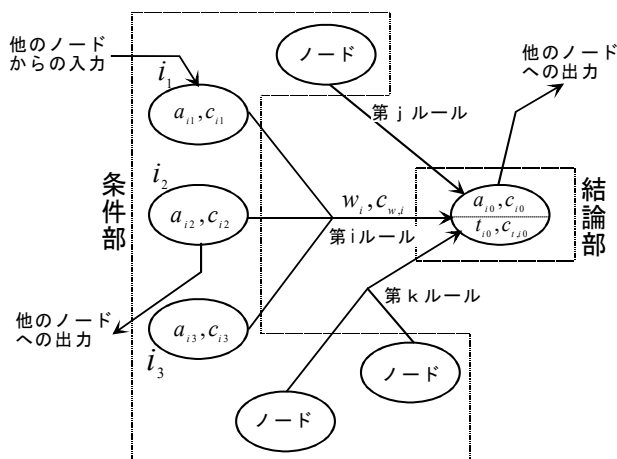


図-1 ルールベース洗練機能付汎用型推論システムの構成¹⁰⁾

工学の分野においてなされている。更に、最近では、ラフ集合論による知識獲得・発見¹⁶⁾や、データベースからの知識発見¹⁷⁾ (KDD: Knowledge Discovery in Databases), あるいはデータ発掘 (Data Mining) の考え方が注目を集めている。

そこで、本研究では、必ずしも広範なルールベースに対する適用性が確認されているとは言えない、皆川らによる知識洗練機能付き推論システムを、鋼橋疲労損傷の補修方法選定問題に適用して、ルールベース洗練機能及び推論機能が対象ドメインによらず有効であることを示す。次に、知識獲得の初期におけるルールベース構築支援を目的として、データベースからの知識発見で用いられる平均圧縮情報量の概念を用いて、データベースからルールを自動生成するアルゴリズムを提案し、その有効性を検証する。

2. ルールベース洗練機能付汎用型推論システムの概要¹⁰⁾

皆川らによる推論システムの全体構成を図-1に示す。同推論システムは、仮説を表すノードと仮説間の関係を表すリンクを推論システムの構成要素とする相互結合ネットワークで表現されている。そして、min-max 演算と、ニューラルネットワークで用いられる誤差逆伝播アルゴリズムを応用した推論方法を用いて、ルールベースの洗練と、通常の仮説推論を行うことができる、ルールベース洗練機能付汎用型推論システムとなっている。

仮説と仮説間の関係において、仮説の成立する可能性と仮説間の関係の強さをそれぞれノード及びノード間の関係とし、これらの属性値としてノード値と結合係数を区間[0, 1]の実数値でそれぞれ表現している。更に各属性値に対して確信度を区間[0, 1]で付与している。また、

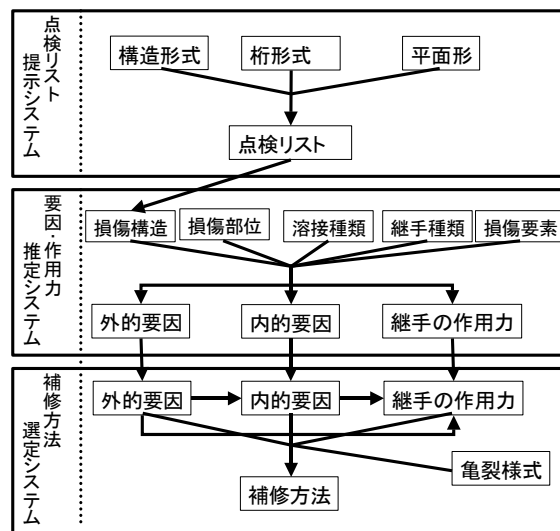


図-2 補修方法選定システムの概要¹⁸⁾

各ノードは過去の事例に関する推論結果を教師データとして持つことができる。以上より、仮説を表すノードは[ノード値, ノード確信度]とこれに対応する[ノード教師値, ノード教師確信度]を属性値として持ち、仮説間の関係を表すノード間の結合は[結合係数, ルール確信度]の属性値を持つ。このように確信度を付与することで情報の不確かさを表現すると共に、仮説推論機能及びルールベース洗練の機能を制御するシステムとなっている。

3. 鋼橋疲労損傷の補修方法選定のためのルールベース

本研究では、田中らにより開発された鋼道路橋に発生した疲労損傷の補修・補強方法選定を対象問題としたエキスパートシステム¹⁹⁾で用いられたルールベースを本推論システムに適用した。

(1) ネットワークの構成

田中らによるシステムでは、図-2に示すように、点検リスト提示システムで、構造形式、桁形式、平面形から損傷事例に見られる損傷構造を探索し、点検すべき構造部分を点検リストとして提示し、このリストを基に点検、調査を行い、損傷が発見されれば、損傷構造、損傷部分、溶接種類、継手種類、損傷要素から要因・作用力推定システムにより、外的要因、内的要因、継手の作用力の推定を行っている。そして、補修の必要性があると判断された場合に、補修方法選定システムにより、要因・作用力推定システムで推定した外的要因、内的要因、継手の作用力の情報に加え観測された事実である亀裂様式の情報を用いて、補修方法の選定を行っている。

しかし、点検リスト提示システム及び、要因・作用力

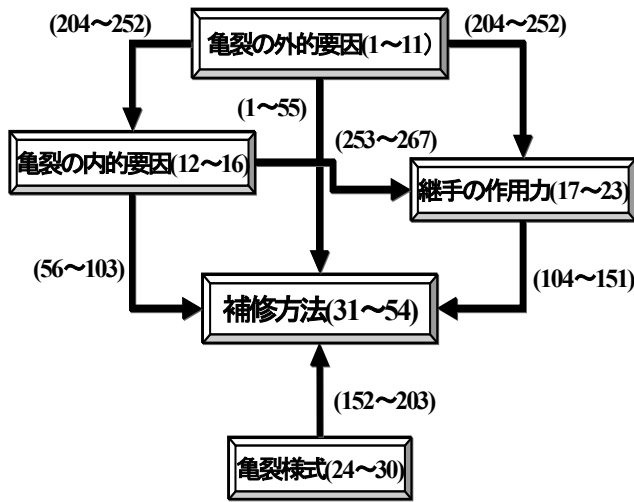


図-3 ネットワークの構成¹⁸⁾

表-1 補修方法項目¹⁸⁾

31	ストップホール
32	ガウジング
33	グラインダ
34	ビーニング
35	腹板ギャップの増大
36	腹板厚の増大
37	切り抜き
38	再溶接
39	フランジと補剛材の溶接
40	再溶融
41	補剛材付き添接板
42	高力ボルト
43	添接板
44	挿入板
45	ケーブルによる吊材相互の緊結
46	横桁と主桁のフランジ連結
47	対傾構と主桁のフランジ連結
48	ダイヤフラムと主桁のフランジ連結
49	横桁とアーチリブのフランジ連結
50	縦桁支承の交換
51	主桁の交換
52	連結板の交換
53	新補剛材の設置
54	制振装置の設置

表-2 提示する事例¹⁸⁾

橋梁名	事例 1	事例 2
国名	米国	米国
架設年度	1958年	1973年
損傷発見年度	12年	0年
構造形式 桁形式 平面形	単純橋 格子合成プレートガーター 直橋	連続橋 合成プレートガーター 直橋
損傷構造 損傷部分	主桁のカバープレートと取り付け部 主桁下フランジ カバープレート	主桁の垂直補剛材取り付け部 主桁腹板 垂直補剛材
溶接種類 継手種類 損傷要素	隅肉溶接 重ね継手 隅肉溶接	隅肉溶接 T継手 主桁腹板
外的要因 内的要因 継手の作用力 亀裂様式 亀裂方向	活荷重の作用 応力集中 重ね継手の作用力12 重ね継手の亀裂様式(i) 溶接に平行な亀裂	輸送架設荷重 2次応力 T継手の作用力3 T継手の亀裂様式(d) 溶接に平行な亀裂
補修方法	ビーニング 再溶融 高力ボルト 添接板	ストップホール グラインダ
損傷状況		

推定システムでは、フレームを用いた推論を行っており、推論過程が明らかにされていない。そこで、亀裂の外的要因、亀裂の内的要因、継手の作用力を亀裂様式と同様、観測された事実である既知情報とみなし、仮説間の関係を相互結合させて、システム内のネットワークを図-3 のように構成した。

(2) 仮説(ノード)の設定

本推論で設定したノードの総数は54個である。このうち、観測された事実の入力情報として30個の入力項目(図-3中のノード番号1~30)を用意し、推論の対象として表-1に示す24個の出力項目(同ノード番号31~54)を用意した。

(3) ルールの提示

田中らによるシステムのルールベースでは、各ルールの因果関係の強さを Necessity, High Possibility, Possibility, Low Possibility の4段階に分類していること

から、本システムに適用するにあたり、結合係数0.8, 0.6, 0.4, 0.2をそれぞれ対応させ、同じく4段階に分類してルールベースの初期状態を構築した。なお、ルールベースのルールの総数は、267個である。

4. 推論システムの適用結果及び考察

上記ルールベースを本推論システムに適用するにあたっては、既知情報とみなした亀裂の外的要因・内的要因、継手の作用力及び亀裂様式に対しては、ノード値として1.0(真の場合)または0.0(偽の場合)を、確信度として1.0を付与した。また、未知情報である補修方法に対しては、真偽が不明であるという意味から、ノード値0.5及び確信度0.1を付与した。

はじめに、表-2に示した事例について、田中らによるシステムの推論結果をルールベース洗練の際に提示する事例として用い、ルール洗練を行った。次に、事例中の補修方法が不明であるとして、仮説推論により補修方法

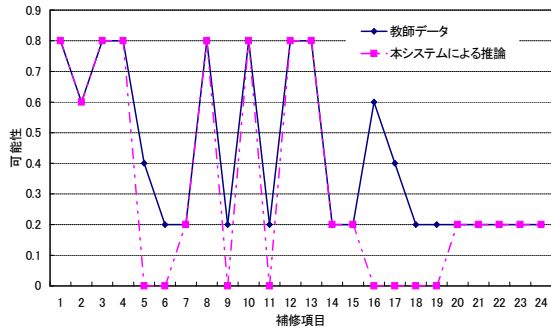


図-4 補修方法の推論結果(事例 1)

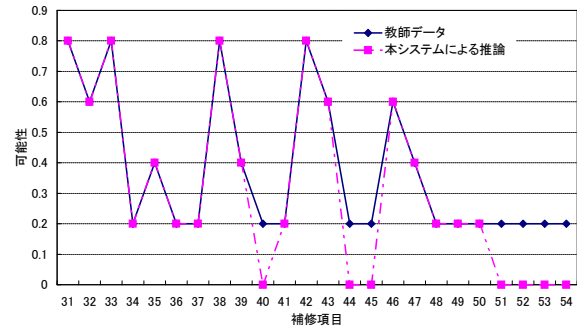


図-5 補修方法の推論結果(事例 2)

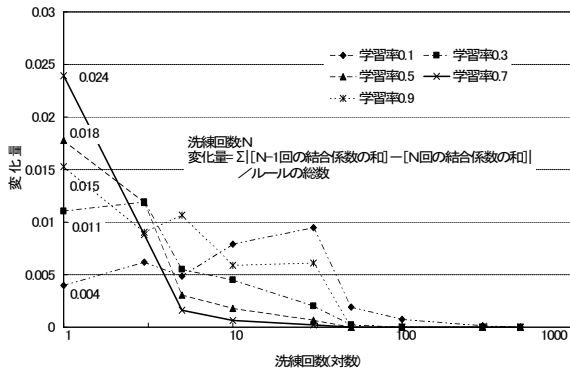


図-6 結合係数の変化量(事例 1)

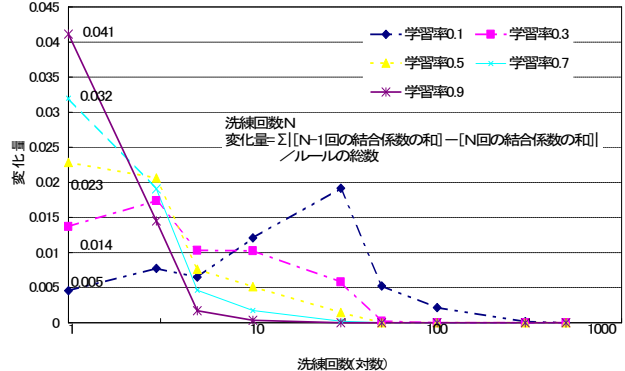


図-7 結合係数の変化量(事例 2)

の選定を行った。図-4 及び図-5 には各事例についての推論結果を示す。これらの図は、横軸に表-1 の補修方法項目のノード番号を、縦軸に補修方法の可能性であるノード値をとり、学習率を 0.3 としたときの推論結果である。また、比較のため、図中には教師データを併せて示した。

図-6,7 には、図中に定義した式より、学習率を変化させた場合の洗練回数毎の結合係数の変化量を求め、横軸にルールベースの洗練回数を対数でとり、縦軸に変化量をとった結果をそれぞれ事例 1,2 について示す。

この結果、補修方法の選定では、実際に取られた補修方法であるストップホール、グラインダが最も高い可能性として推論されており、推論全体の傾向も教師データと同様の結果を示していることから、推論システムとしての機能は、十分果たしていると考えられる。また、ルールベースの洗練状況も学習率の違いにより変化の傾向は違うが、洗練回数の増加と共に収束しており、これに加え推論結果が妥当な解を示していることから、本システムのルール洗練機能も十分機能しているといえる。

これらの結果より、本システムが鋼橋疲労損傷の補修方法選定問題においても、本来エキスパートシステムに要求される対象問題の可能性の分類を行い、技術者支援システムとして機能していることが示された。また、本システムのルール洗練機能が、目標とする知識の共有・再利用といった知識獲得作業に対し、有効に機能していることも示された。

5. 平均圧縮情報量によるデータベースからの知識発見

前節では、皆川らが提案した相互結合型のネットワークで構成された推論システムの知識洗練機能が対象ドメインによらず有効であることを示した。この方法では、ある程度の妥当性の確認された推論システムと知識ベース、あるいは事例ベースを用いることにより、それらと適合するような知識ベースを獲得する。したがって、前者の場合には既存の知識ベースの枠内での知識洗練であるし、後者の場合には、既存の事例ベースの枠内でのルール獲得となる。本節では、これらの枠組みを越えて新たな知識を発見するための方法論として、情報理論の分野で用いられている平均圧縮情報量の知識発見への適用性を検討する。対象ドメインは前節と同様である。

(1) 平均圧縮情報量 (Average Compressed Entropies)

データベース中には、有用性の有無に関わらず、極めて多くの知識が存在する。この中から有用である知識、つまり興味深い知識を見分けることは非常に困難であり、これを見分けるには、対象領域の背景知識や、対象領域固有の規準に基づいた興味深さの評価を行うのが通常である。例えば、皆川らのルールベース洗練機能付推論システムでのルールの結合係数や確信度の付与の方法が、専門家による主観性に依存して行われていることも一例

であるといえる。

一方、抽出した知識、あるいはルールは、このような背景知識によらず、客観的、定量的に評価されることが望ましい。そこで、情報理論的観点から、データベースよりルールを抽出するために圧縮されるデータの情報量である平均圧縮情報量を求める。これにより、抽出したルールの持つ興味深さを評価することとした。

(2) データベースにおける事例とルール

データベース中に存在する各事例を e_i とし、事例の総数を N とする。つまり、データベース中には、以下のように N 個の事例が存在する。

$$\{ N \cdot e_1, e_2, \dots, e_i, \dots, e_N \}$$

また、事例 e_i は、 m 個の要素 a_{im} によって構成されている。よって事例 e_i は、次のように表すことができる。

$$\{ e_i : a_{i1}, a_{i2}, \dots, a_{im} \}$$

データベース内から、ルールを生成する際に、条件部になり得る要素 a_i を y 、結論部となり得る要素 a_m を x とし、条件部と結論部になる要素を区別して表すことにする。なお、各要素は、条件部の要素にも結論部の要素にもなり得るものとする。

次に、データベース内の各事例 e_i において、条件部の要素 y を持つ事例の数を n_y とし、結論部の要素 x を持つ事例の数を n_x とする。また、条件部の要素 y と結論部の要素 x を共に持つ、すなわち、ルール $y \rightarrow x$ を抽出することのできる事例の数を $n_{y,x}$ とする。

(3) 平均圧縮情報量による評価アルゴリズム

はじめに、結論部の要素 x の生起確率 $p(x)$ を式(1)より求め、結論部の要素が x でない場合、 $p(\bar{x})$ の生起確率を式(2)より求める。

$$p(x) = \frac{n_x}{N} \quad (1)$$

$$p(\bar{x}) = 1 - \frac{n_x}{N} \quad (2)$$

次に、結論部の要素 x と条件部の要素 y の結合確率 $p(x, y)$ を式(3)より求め、また、結論部の要素が \bar{x} で、条件部の要素が y の結合確率 $p(\bar{x}, y)$ を式(4)より求める。

$$p(x, y) = \frac{n_{y,x}}{N} \quad (3)$$

$$p(\bar{x}, y) = 1 - \frac{n_{y,x}}{N} \quad (4)$$

ここで、結合確率は、データベース内の事例 N 個のう

ち、ある事例 e_i について、条件部の要素 x と結論部の要素 y を同時に持つ事例の割合を示す。

また、条件部の要素 y が起こった時に結論部の要素 x が起こる、結論部の要素 x についての条件付確率 $p(x|y)$ を式(5)より求め、条件部の要素 y が起こった時に、結論部が \bar{x} である条件付確率 $p(\bar{x}|y)$ を式(6)より求める。

$$p(x|y) = \frac{n_{y,x}}{n_y} \quad (5)$$

$$p(\bar{x}|y) = 1 - \frac{n_{y,x}}{n_y} \quad (6)$$

情報理論的には、通常、結論部の要素 x がデータベース内で持つ情報量を、 $-\log\{p(x)\}$ で示す。また、この時、結論部の要素 x を事例内の要素 $x = a_m$ として持つ事例が、同時に条件部の要素 y を $y = a_i$ として持つ場合があった時、 $y \rightarrow x$ というルールが事例中に内在していることになる。この要素 y と要素 x を同時に持つ時の結論部の要素 x がデータベース内で持つ情報量は、 $-\log\{p(x|y)\}$ で示すことができる。この2つの情報量により、事例中に内在する $y \rightarrow x$ というルールの情報量は、条件部の要素 y が起こる前の情報量 $-\log\{p(x)\}$ と、条件部の要素 y が起こった後の情報量 $-\log\{p(x|y)\}$ との差によって表すことができる。つまり、ルール $y \rightarrow x$ を生成することにより、結論部の要素 x についての情報量を圧縮したことになる。

また、式(3)に示したように、ルール $y \rightarrow x$ を生成することが可能な事例は、データベース内に $N \cdot p(x, y)$ 個存在していることから、データベース内において結論部の要素 x が、条件部の要素 y によって圧縮される情報量は、 $N \cdot p(x, y)$ と各情報量との積により求めることができる。以上に示したことは、同様に結論部の要素が \bar{x} でない場合、 $p(\bar{x}, y)$ についてもいえる。

この $N \cdot p(x, y)$ あるいは $N \cdot p(\bar{x}, y)$ と、情報量の積を圧縮情報量 (Compressed Entropies : CE) と呼び、式(7)により $CE(x, y)$ で表す。

(7)

この圧縮された情報量を評価値とするため、データベース内の事例の総数 N により平均値をとって、評価値となる平均圧縮情報量 $ACE(x, y)$ を式(8)より求める。

$$\begin{aligned} ACE(x, y) &= CE(x, y) / N \\ &= p(x, y) \log \frac{p(x|y)}{p(x)} + p(\bar{x}, y) \log \frac{p(\bar{x}, y)}{p(\bar{x})} \end{aligned} \quad (8)$$

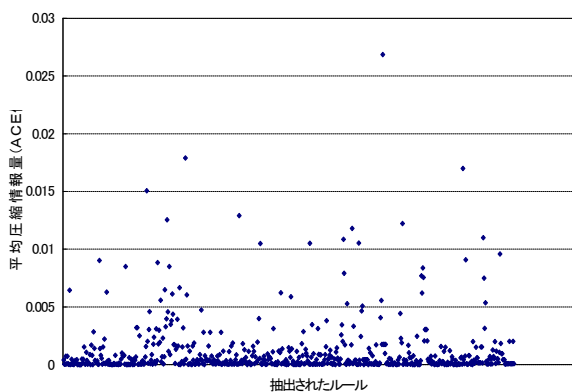


図-8 ACE 値により抽出されたルール

6. 適用結果及び考察

前述の平均圧縮情報量に基づく知識発見の適用例として、前節と同様、鋼橋疲労損傷の補修方法選定問題を選び、既存の推論システムで得られる事例群を生データのデータベースと見なし、ここからルールを抽出した後、それをルールベースとする推論を実施して、抽出されたルールの有効性を検討する。

(1) データベースの作成

ここでは、平均圧縮情報量によって、データベースから有効な知識を抽出することが可能であるかを検討するため、田中らによる鋼橋疲労損傷の補修方法選定システムのルールベースを基に、総数 2695 の仮想事例を作成し、これをデータベースとした。

仮想事例の作成は、亀裂の内的要因、外的要因、継手の作用力、亀裂様式の 30 個の入力情報項目について、起こり得る可能性のある全組み合わせ、2695 通りを仮に発生した事実とし、これを仮想事実とした。この仮想事実に対して、皆川らのルールベース洗練機能付推論システムより、補修方法を唯一選定し、仮想事実と組み合わせで仮想事例とした。

(2) データベースからのルール抽出

平均圧縮情報量である式(8)を用いたルール抽出アルゴリズムを仮想事例より作成したデータベースに適用してルール抽出を試みた。その結果、620 のルールがデータベースから抽出された。図-8 には、横軸に抽出された 620 のルールをとり、縦軸に評価値である平均圧縮情報量 (ACE 値) をとった図を示す。

この抽出されたルールを基に新たにルールベースを構築し、鋼橋疲労損傷の補修方法の選定を行って、抽出されたルールの有用性を検討した。

抽出されたルールのうち、ACE 値の高いルールから順に、次に示す 3 通りの規準で結合係数を付与し、それぞれ新

表-3 ACE 値を基に構築したルールベース (b の場合)

	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
1	0.4	0.2			0.6			0.2	0.6			0.2			0.8									
2		0.2	0.6					0.4				0.4	0.4											
3	0.2	0.4										0.8		0.6			0.4					0.6	0.8	
4	0.2	0.2	0.6	0.4				0.2	0.2			0.6										0.8	0.6	
5		0.8	0.6	0.4	0.4			0.4	0.4															
6	0.6	0.6	0.6	0.2	0.6							0.2	0.2											
7	0.4	0.2	0.4																			0.2		
8		0.6						0.4					0.2	0.6								0.6	0.6	
9	0.6	0.6		0.8									0.6											
10			0.6	0.4				0.6				0.4	0.6											0.8
11		0.2	0.8	0.2	0.4			0.2	0.4			0.4	0.6											0.8
12	0.4							0.6				0.8	0.6											
13	0.8	0.4	0.8					0.6				0.8	0.6											
14	0.6	0.2	0.8					0.8				0.4	0.2											
15	0.8	0.4	0.8	0.8	0.8			0.8	0.8			0.4	0.6											
16	0.2	0.6	0.8	0.8	0.8			0.8	0.8			0.6	0.8	0.6	0.8							0.8	0.8	0.8
17	0.2	0.8	0.8	0.8	0.8			0.8	0.8			0.2												
18	0.8	0.8	0.2	0.4	0.8			0.8	0.8			0.8												
19	0.6	0.2		0.8	0.6			0.4	0.8			0.2	0.4									0.6	0.4	0.2
20	0.6	0.6	0.8	0.4				0.2	0.8			0.6												
21	0.4	0.6	0.4	0.4	0.8			0.4	0.8			0.4	0.6	0.4			0.2					0.6	0.4	0.2
22	0.4	0.6	0.4	0.2	0.6			0.2	0.8			0.2	0.6	0.4			0.2					0.4	0.4	0.2
23	0.2	0.6	0.2	0.4	0.8			0.2	0.8			0.2	0.8				0.6					0.6	0.2	0.2
24	0.8	0.2	0.4	0.4	0.2	0.6		0.8	0.6			0.6	0.4				0.6							
25		0.4	0.4		0.6			0.6	0.6			0.6	0.4				0.6							
26	0.6	0.8	0.8	0.6	0.2			0.6	0.4			0.4	0.8				0.6							
27	0.8	0.6	0.8	0.2				0.2	0.6			0.6	0.4				0.6							
28		0.2	0.6	0.6				0.6	0.6			0.6	0.2				0.2							
29	0.6		0.8	0.2				0.4	0.2			0.4	0.6	0.8			0.6					0.8	0.2	0.8
30	0.4	0.2	0.4	0.2	0.6			0.4				0.6	0.6		0.8		0.6					0.8	0.8	0.2

たなルールベースを構築した。

- a) 全ルールを 5 分割し、上位 40% を 0.7、次の 40% を 0.4 とした場合
- b) 全ルールを 5 分割し、上位から 20% 毎に 0.8、0.6、0.4、0.2 とした場合
- c) 全ルールを 10 分割し、10% 毎に、0.8 から 0.1 までとした場合

なお、3 通りすべてにおいて、下位 20% のルールについては切り捨てとした。

表-3 には、b) の場合のルールベースを示す。表中の番号 1~30 は、図-3 の入力情報項目の番号に対応しており、また、31~54 は、図-3、及び表-1 の補修方法項目の番号に対応している。

(3) 補修方法選定結果

はじめに、鋼橋疲労損傷の補修方法選定を対象に、新たに構築したルールベースを適用して、補修方法の選定を行った。これにより、抽出されたルールを用いて構築したルールベースと、田中らによるルールベースをそれぞれ適用した場合での補修方法選定の結果を比較することで、抽出されたルールの妥当性を検討することができる。推論に適用した事例は、表-2 で提示した事例 1, 2 である。

図-9、図-10 に a) の場合における補修方法選定結果をそれぞれ事例 1、事例 2 について、横軸に補修方法項目の番号をとり、縦軸に各項目の可能性であるノード値をとって示す。比較のため、田中らによるルールベースで推論した推論結果、及び教師データを併せて示す。以下、図-11、図-12 には b) の場合の結果を示し、図-13、図-14 には c) の場合の結果を示す。

この結果、a) の場合、図-8 のように、ACE 値による評価によって幅広く分布するルールを 0.7 と 0.4 の 2 段階で分類したため、ノード値による可能性の分類が十分になされていないことがわかる。これに対し、b)、c) の場合は、それぞれ 5 段階、10 段階に分類したことから、

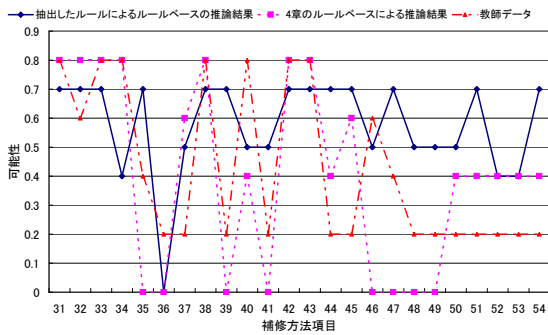


図-9 a)の場合の補修方法選定結果(事例1)

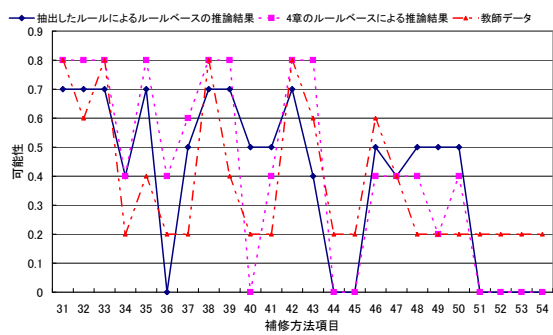


図-10 a)の場合の補修方法選定結果(事例2)

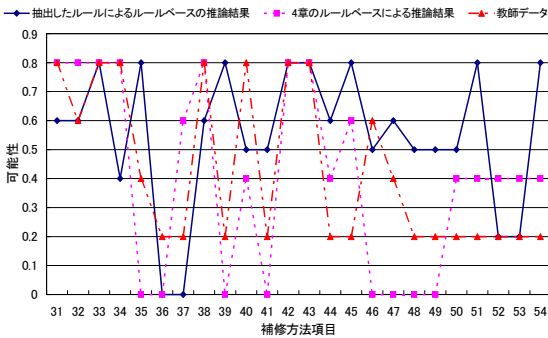


図-11 b)の場合の補修方法選定結果(事例1)

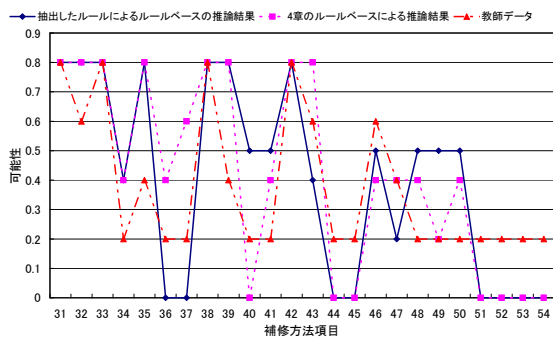


図-12 b)の場合の補修方法選定結果(事例2)

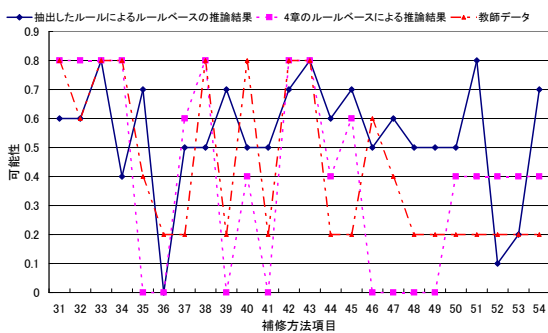


図-13 c)の場合の補修方法選定結果(事例1)

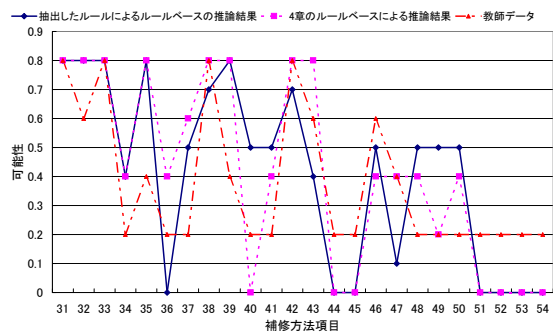


図-14 c)の場合の補修方法選定結果(事例2)

共に十分な可能性の分類を行って推論結果を提示している。図-11、図-13の事例1の結果では、教師データとの間に大きな違いが見られる。特に、補修方法項目の番号51、54で、田中らのルールベースでの推論結果、及び教師データ共に0.4、0.2と低い可能性を示しているのに対し、0.8と高い可能性を示しているなど、ノード値が0.8を示す項目についての違いが目立つ。しかし、このことは、田中らによるルールベースを用いた場合の推論結果についても同様である。また、図-12、図-14の事例2の結果では、ノード値0.8を示す補修方法項目は、ほぼ教師データと一致しており、b)の場合で比較すると、田中らのルールベースによる推論結果よりも精度の良い結果を示している。また、推論結果全体の傾向としても、可能性の値に違いは見られるが、田中らによるルールベースでの推論結果、及び教師データの傾向と同様な結果を

示している。

しかし、c)の場合、結合係数を細分化したことにより、ノード値0.8を示す可能性の高い項目の提示が、特に事例1については3項目と少数であることから、本来システムに要求されている技術者支援システムとしての可能性の分類の観点からすると、利用者が提示された解を参照する以前に、過度に可能性が分類される恐れがある。

(4) 抽出されたルールと既存ルールと比較

ここでは、(3)において事例1と事例2の間で推論結果にある程度の差が生じていることから、この原因を検証するため、図-3に示した各入力情報と補修方法の結合に対して、抽出されたルールと既存ルールとを比較する。

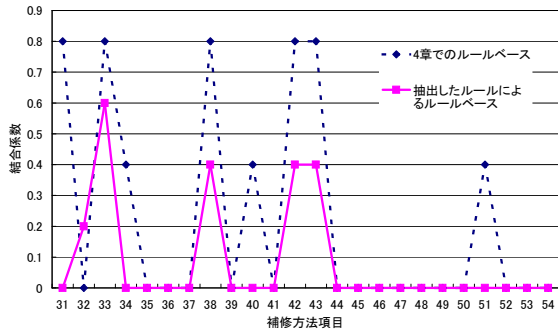


図-15 外的要因 (2) と補修方法の因果関係 (事例 1)

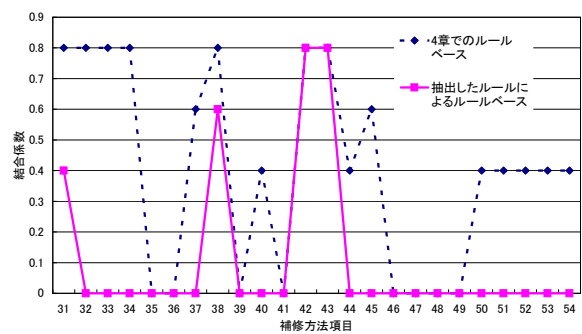


図-16 内的要因 (12) と補修方法の因果関係 (事例 1)

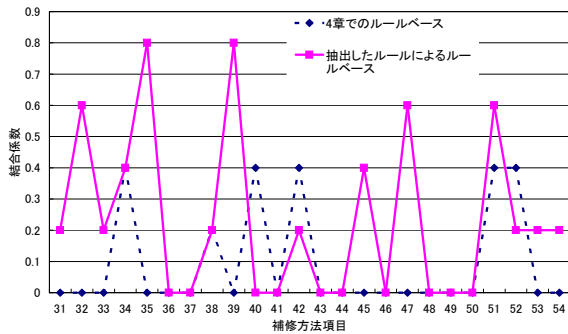


図-17 継手の作用力 (23) と補修方法の因果関係 (事例 1)

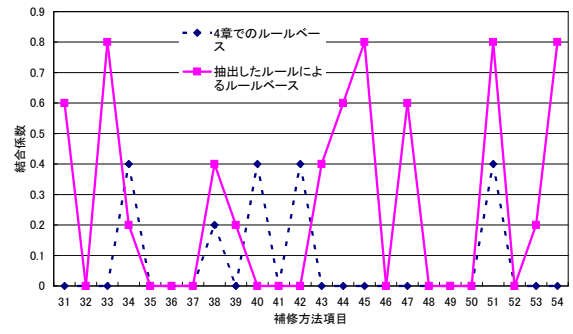


図-18 亀裂様式 (29) と補修方法の因果関係 (事例 1)

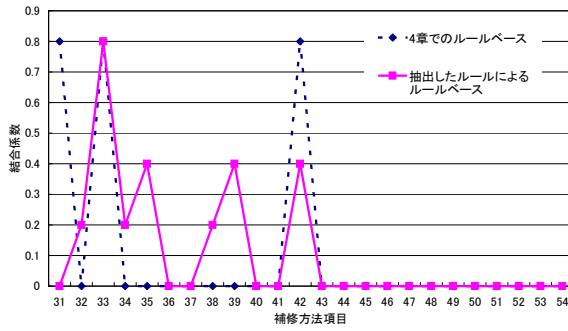


図-19 外的要因 (11) と補修方法の因果関係 (事例 2)

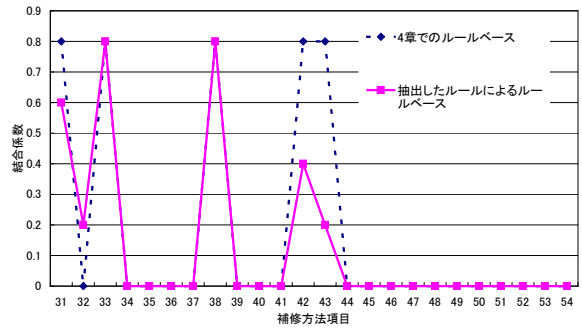


図-20 内的要因 (14) と補修方法の因果関係 (事例 2)

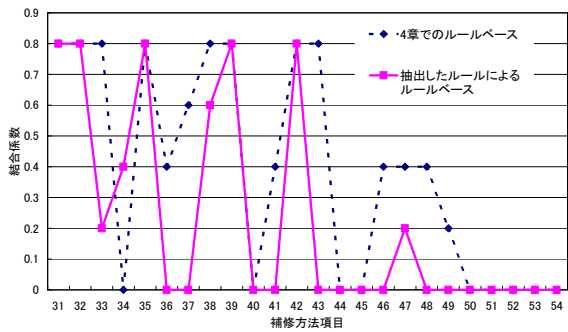


図-21 継手の作用力 (18) と補修方法の因果関係 (事例 2)

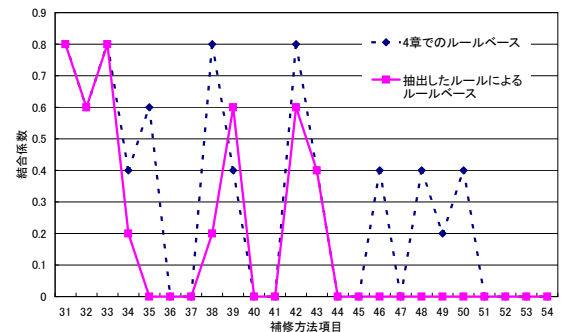


図-22 亀裂様式 (27) と補修方法の因果関係 (事例 2)

事例1については図-15から図-18に、事例2については図-19から図-22に、それぞれ入力情報項目と補修方法項目との結合係数を、横軸に補修方法項目(番号34-54)をとり、縦軸に各項目の結合係数をとって示す。なお、図-15から図-22は、それぞれ以下に示す入力情報についての補修方法との結合関係を示した図である。

図-15：事例1の亀裂の外的要因〔活荷重の作用(入力情報項目2)〕

図-16：事例1の亀裂の内的要因〔応力集中(入力情報項目12)〕

図-17：事例1の継手の作用力〔重ね継手の作用力②(入力情報項目23)〕

図-18：事例1の亀裂様式〔重ね継手の亀裂様式(i)(入力情報項目29)〕

図-19：事例2の亀裂の外的要因〔輸送架設荷重(入力情報項目11)〕

図-20：事例2の亀裂の内的要因〔2次応力(入力情報項目14)〕

図-21：事例2の継手の作用力〔T継手の作用力③(入力情報項目18)〕

図-22：事例2の亀裂様式〔T継手の亀裂様式d(入力情報項目27)〕

この結果、事例1では、図-11において補修方法項目51, 54がノード値0.8を示しているのは、図-18の亀裂様式と補修方法との関係で、結合係数が、田中らのルールベースでは補修方法項目51に弱い因果関係を持たせているだけであるのに対し、抽出したルールでは、51, 54共に強い因果関係を持っていることが示されている。このように、事例1では、田中らのルールベースでは弱い因果関係であったり、全く因果関係を持たないルールであるのに対し、ルール抽出アルゴリズムでは強い因果関係を持ってルールを抽出していることが、図-17や図-18の他の項目に関しても示されており、この結果、事例1の推論では、田中らのルールベースでの推論結果、及び教師データの値との間で違いが生じたと考えられる。

図-19から図-22より、事例2では、田中らのルールベースで強い因果関係を持つルールのほとんどに関して、同様に強い因果関係を持ってルールを抽出していることがわかる。そのため、事例2の推論では、図-21において低い可能性のノード値しか示していない補修方法項目番号36, 37において、田中らのルールベースでは、関係は弱い因果関係を持っているルールに対し、ルール抽出アルゴリズムではルールを抽出していない程度であり、推論全体の傾向も4.での推論結果、及び教師データと同様の結果を示している。

7. おわりに

本論では、はじめに知識の共有・再利用を主な目的としたルールベース洗練機能付推論システムが、対象問題によることなく機能するかを検討した。この結果、仮説推論の推論結果、及びルールベースの洗練による結合係数の変化状況から、対象問題を鋼橋疲労損傷の補修方法選定とした場合にも、推論システムがエキスパートシステムとしての機能を十分果たしていることが示された。

次に、知識獲得の初期における知識獲得問題解決のため、データベースからの知識発見(KDD)の手法に平均圧縮情報量(ACE値)を用いてルール抽出を試みた。また、抽出したルールの有用性を検討するため、抽出したルールを用いて鋼橋疲労損傷の補修方法選定のためのルールベースを構築し、実際の事例を適用して補修方法選定を試みた。この結果、推論結果は、田中らによるルールベースでの推論結果と同等の結果を示していることから、知識発見におけるルール抽出にACE値が十分有効であると考えられる。

謝辞：本研究を行うにあたり、関西大学の田中成典助教授に有益なご助言を頂きました。また、本研究室の4年生、特に稲蔵君、及び関田君の多大な協力に感謝いたします。

参考文献

- 1) 本位田 真一, 市川 照久: エキスパートシステム基礎技術, オーム社, 1989.12.
- 2) S. M. ワイス, C. A. クリコフスキ(森 健一訳): エキスパートシステムの設計, 近代科学社, 1987. 7.
- 3) 宮本 文穂, 森川 英典, 益成 一郎, 古川 正典: 橋梁診断エキスパートシステムにおける知識の更新手法, 構造工学論文集, Vol.37A, pp.643-654, 1991.3.
- 4) 串田 守可, 徳山 貴信, 宮本 文穂: 橋梁診断における経験的知識に内在するあいまいさの定量化に対するファジィエントロピーの適用, 構造工学論文集, Vol.38A, pp.571-584, 1991.3.
- 5) 三上 市蔵, 田中 成典, 倉地 晶: 鋼橋疲労損傷の補修方法選定のための学習機能のあるニューラルネットワークシステム, 構造工学論文集, Vol.37A, pp.655-668, 1991.3.
- 6) 三上 市蔵, 田中 成典, 倉地 晶, 米田 慎二: 鋼橋疲労損傷の補修方法選定システムにおける類推推論と負の学習の実現, 構造工学論文集, Vol.38A, pp.557-569, 1992.3.
- 7) 田中 成典, 三上 市蔵, 前田 秀典: 事例ベース推論を用いた鋼道路橋疲労損傷の補修方法の選定システム, 第2回ファジィ土木応用シンポジウム講演論文集,

- pp.37-42, 1994.12.
- 8) 田中 成典, 三上 市蔵, 前田 秀典: 動的矛盾知識の検出・解消処理を施した鋼橋疲労損傷補修方法選定システム, 構造工学論文集, Vol.41A, pp.587-596, 1995.3.
 - 9) 田中 成典, 三上 市蔵, 前田 秀典, 小林 篤司: 鋼道路橋の疲労亀裂に対する補修・補強方法の類推推論選定システム, 構造工学論文集, Vol.42A, 1996.3.
 - 10) 皆川 勝, 佐藤 茂, 上谷 丈和: 事例ベース推論を援用した知識洗練機能付診断エキスパートシステムの開発, 土木学会論文集(投稿中).
 - 11) 田淵 真, 田浦 俊春: 遺伝的学習機構と人との対話型知識獲得手法, 人工知能学会誌, Vol. 11No. 4 pp. 600-607, 1996. 7.
 - 12) 吉田 健一, 元田 浩: 逐次ペア拡張に基づく機能推論, 人工知能学会誌, Vol. 12No. 1, pp. 58-67, 1997. 1.
 - 13) 河野 毅, 濱田 進, 荒木 大, 小島 昌一, 田中 利一: 事例ベース推論によるエラー補修と知識獲得人工知能学会誌, Vol. 9No. 3, pp. 408-416, 1994. 5.
 - 14) 樽松 理樹, 山口 高平: 事例に基づく推論とモデル推論の統合に基づく知識獲得支援システム(1), 人工知能学会誌, Vol. 11No. 4, pp. 585-592, 1996. 7.
 - 15) 山口 高平, 樽松 理樹, 下津 直武, 中尾 博司, 落水 浩一郎: 事例に基づく推論とモデル推論の統合に基づく知識獲得支援システム(2), 人工知能学会誌, Vol. 11No. 4, pp. 593-599, 1996. 7.
 - 16) 中村 昭, 津本 周作, 田中 博, 小林 聡: ラフ集合とその応用, 人工知能学会誌, Vol. 11. No. 4, pp. 209-215, 1996. 5.
 - 17) Bing Leng and Bruce G. Buchanan: Using knowledge-assisted discriminant analysis to generate new comparative terms; Artificial Intelligence and Statistics IV, Springer-Verlag, pp.479-487, 1993.
 - 18) 田中 成典: 橋梁工学への知識情報処理技術の応用に関する研究, 関西大学学位論文, pp.25-248, 1996.9.

KNOWLEDGE DISCOVERY WITH AVERAGE COMPRESSED ENTROPY FOR SELECTING RETROFITTING METHOD OF STEEL BRIDGES DAMAGED BY FATIGUE

Takekazu KAMITANI supervised by Masaru MINAGAWA

Recently, Minagawa proposed a knowledge refinement method for crack diagnostic expert system. The inference engine was constructed for a reciprocal network based on min-max composition algorithm. The knowledge refinement function was installed into the engine by using the concept of back propagation algorithm. First the inference and refinement method are applied to the rule-base system for selecting the retrofitting method of steel bridges damaged by fatigue. It is confirmed that the inference engine can be used for any particular domain. Second, for the purpose of knowledge discovery, we evaluated average compressed entropies for a case-base virtually constructed through some inferences with the inference system that we proposed for selecting the retrofitting method. It is found from the analyses that the average compressed entropy is an effective measure for the discovery of knowledge that is implicitly buried into data-bases or case-bases.