

# ラフ集合の概念による橋梁伸縮継手損傷 のルール型知識獲得

加賀山泰一<sup>1</sup>・河村 圭<sup>2</sup>・宮本文穂<sup>3</sup>・田中信也<sup>4</sup>

<sup>1</sup>正会員 工修 阪神高速道路公団 工務部(〒541-0056 大阪市中央区久太郎町4-1-3)

<sup>2</sup>正会員 博士(工学) 山口大学 工学部知能情報システム工学科(〒755-8611宇部市常盤台2-16-1)

<sup>3</sup>フェロー会員 工博 山口大学教授 工学部知能情報システム工学科(〒755-8611宇部市常盤台2-16-1)

<sup>4</sup>修士(工学) ㈱宇部情報システム(〒755-8622宇部市相生町8-1宇部興産ビル)

本研究では、橋梁の合理的な維持管理手法を立案する一環として、データベースに蓄積されている情報から、データマイニングの手法を用いて、橋梁伸縮継手の損傷に関するルール型知識の獲得を試みた。知識獲得のためのデータマイニングの手法としては、ラフ集合の概念を取り入れた決定表簡約化を採用した。既存の研究において、このような知識を抽出したという研究はいくつか報告されているが、獲得されたルールの評価までを行った研究は数少ない。そこで、本研究では、橋梁伸縮継手の損傷に関するルール型知識を抽出するだけでなく、獲得された知識を評価するための方法を提案し、その検証を行った。

*Key Words* : data mining, rough sets, data base, expansion joint, knowledge discovery,

## 1. はじめに

近年、情報処理技術の発展やコンピュータ技術の急速な進展に伴い、データベースに関する技術であるデータマイニングが注目を集めている。データマイニングとは、膨大なデータの山には、表面には現れていない隠れた有用な情報が多数存在している可能性があり、それらの情報を掘り起こすという考え方である<sup>1)~7)</sup>。データマイニングの問題としては、類似性からデータをグループ分けする分類、関連性の強いデータの組合せを発見する連関、既知データから規則性を発見する判別、既知データから将来の傾向を見つける予測といったものがあり、手法としては、クラスター分析、主成分分析、時系列パターン分析、相関分析、決定木、回帰分析、ニューラルネットワークなどがある。最近では、金融、通信、流通などの不特定多数の顧客を相手にするビジネス分野において、マーケティング活動を支援する目的で浸透している<sup>8),9)</sup>。すなわち、顧客に関する情報をデータベース化し、時系列的に蓄積されたデータを分析、加工しながら何らかの仮説を導き出し、新たなマーケティング施策として顧客にフィードバックするというデータベースマーケティングである。

橋梁構造物の分野においても、維持管理に対する認識の高まりから、各機関において橋梁台帳等のデータベース化が急速に行われている。しかしながら、これらのデ

ータベースの多くは、単に従来紙面で管理していた情報を電子化し、検索や集計に利用しているだけのシステムが多く、記憶媒体の小型化や検索の迅速化がなされただけで、蓄積されているデータが有効に利用されているとは言い難い。これらのデータを詳細に分析することによって、構造物の劣化進行の予測など、維持管理に関する有用な知見を得ることも可能であると思われるが、対象となるデータ数は膨大であり、分析手法も確立されていないため、単純な統計解析を除けば、体系的な分析が行われていないのが実状である。しかし、過去よりデータベースに蓄積されている膨大なデータには、我々がまだ把握していない橋梁部材の劣化・損傷に関する何らかの有益な情報が隠れている可能性は十分にあり、データマイニングの手法を用いることにより、新たな情報が掘り起こせたならば、それらの情報は、橋梁維持管理を行う際の重要な知識や指標に成り得るとともに、今後の合理的な橋梁の維持管理手法の確立に大きく貢献できるものと考えられる。

以上の背景から、本研究では、橋梁データベースに蓄積されている膨大な維持管理に関するデータの内、伸縮継手の損傷に関するデータを対象として、データマイニング手法によるルール型知識の獲得を試みた。データマイニング手法としては、事例データベースからルール型知識を抽出でき、また、判別と連関といった問題に適用する手法として注目されているラフ集合の概念を利用し

た決定表簡約化手法を採用した。過去の研究において、実際に橋梁データベースから橋梁部材劣化に関する知識を獲得したという研究も報告されているが、獲得された知識の評価までを行った研究は数少ない<sup>10)~13)</sup>。したがって、本研究では、伸縮継手の損傷に関するルール型知識を獲得するだけでなく、獲得された知識の評価方法を提案し、その評価を行った。

## 2. ラフ集合を用いたデータマイニング

### (1) ラフ集合の基本概念

ラフ集合は、1982年にポーランドの Zdzislaw Pawlak により提唱されたあらい (rough) 知識情報を取り扱う集合理論である<sup>14)</sup>。ラフ集合は、対象に関する属性と属性値を用いて、その対象と対象外とを識別する。それにより、対象について識別できるものと識別できないものがあることが明確になる。つまり識別不能性 (indiscernibility) が定義でき、その識別できる範囲とともに判断、処理する方式が得られることとなる。対象物の属性がこの分類分けに対して全て同じであるならば、それらの対象物は識別不可能であり、同じものとして取り扱われる<sup>15)</sup>。

ここで橋梁を具体例にすると、「主桁材料」という属性のみを識別の基準にした場合について説明する。この場合、属性値を「PC桁」とすれば、PC桁であればどのような形式、規模、橋齢(しゅん工年)であろうとも全て同じ橋梁として識別されることとなる。この識別不能性が、ラフ集合の最も基本的な概念である。ここで、各橋梁を詳細に識別するためには、識別の基準である属性の数を増やすことで可能となる。上記の例の場合、「主桁材料」の他に、「床版材料」、「設計荷重」、「橋長」などの属性を加えることにより、より細かな識別ができ、これら属性を無限に増やせば、全ての橋梁の識別が可能となる。

このような識別の性質を利用することにより、データベースの膨大なデータの中から、新たな知識が発見される可能性がある。

### (2) 決定表からの知識の獲得

#### a) 決定表

ここでは、ルール獲得に利用する決定表とその基本的性質について述べる<sup>10)</sup>。いま、次のような情報システム  $S = (U, A)$  を考える。ここで、 $U$  は対象の集合、 $A$  は属性の集合である。 $P, Q (\in A)$  を各々条件属性、決定属性と呼ばれる属性からなる  $A$  の2つの部分集合とする。このように、2つの異なる属性の部分集合を持った

表-1 橋梁データにおける決定表の例

橋梁	条件属性			決定属性
	橋長	主桁材料	設計荷重	桁の損傷判定
$x_1$	10	PC	TL-20	A
$x_2$	15	RC	TL-25	B
$x_3$	10	PC	TL-20	B
$x_4$	15	Steel	TL-20	C
$x_5$	20	PC	TL-20	A
$x_6$	20	RC	TL-20	A
$x_7$	10	Steel	TL-25	C
$x_8$	20	Steel	TL-25	B

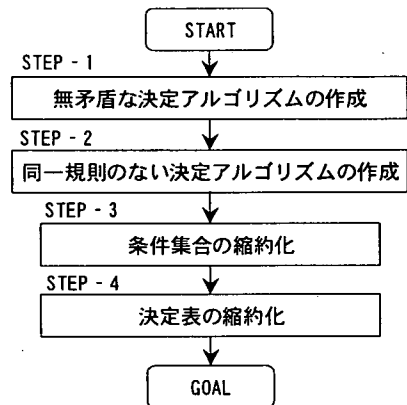


図-1 決定表の簡約化の手順

情報システム  $S = (U, A)$  を決定表と呼び、 $T = (U, A, P, Q)$  と表わす。決定表の具体例として表-1に橋梁データの例を示す。

#### b) 決定表の簡約化

決定表を用いたルール獲得とは、すなわち決定表の簡約化である。決定表に関わる多くの応用領域において、与えられた決定表を簡約化するという問題は、最も重要である。簡約化の一例として、決定表における条件属性の簡約化が挙げられる。条件属性の簡約化とは、全体として決定表の識別能力を損なわずに、ある条件属性を削除するということである。これにより簡約化された決定表では、より少ない条件でもって同一の決定を下すことが可能となる。この種の簡約化によって、不要な条件を調べるといった作業から開放されることとなる。

決定表の簡約化の具体的な手順について図-1に示すとともに、表-2のシステムを例題として、その手法を具体的に説明する<sup>10)</sup>。

表-2 決定表

事例	条件属性				決定属性
	a	B	c	d	e
1	1	1	1	2	1
2	1	2	2	2	1
3	2	1	2	2	1
4	3	1	3	2	1
5	1	1	3	2	2
6	1	2	1	2	2
7	2	1	1	2	2
8	2	2	1	2	2
9	3	2	1	2	2
10	3	2	1	2	2
11	1	1	1	1	2
12	1	1	1	1	3
13	1	1	1	1	3
14	1	1	1	1	3
15	1	1	1	1	3
16	1	1	1	1	3
17	1	1	1	1	3
18	1	1	1	1	3
19	1	1	1	1	3
20	1	1	1	1	3

表-3 矛盾の無い決定表

事例	条件属性				決定属性
	a	B	c	d	e
1	1	1	1	2	1
2	1	2	2	2	1
3	2	1	2	2	1
4	3	1	3	2	1
5	1	1	3	2	2
6	1	2	1	2	2
7	2	1	1	2	2
8	2	2	1	2	2
9	3	2	1	2	2
10	3	2	1	2	2

表-4 矛盾している決定表

事例	条件属性				決定属性
	a	B	c	d	e
11	1	1	1	1	2
12	1	1	1	1	3
13	1	1	1	1	3
14	1	1	1	1	3
15	1	1	1	1	3
16	1	1	1	1	3
17	1	1	1	1	3
18	1	1	1	1	3
19	1	1	1	1	3
20	1	1	1	1	3

[STEP-1] 矛盾のない決定アルゴリズムの作成

最初に与えられた決定表(決定アルゴリズム)に矛盾がある場合、矛盾の無い決定表を作成する必要がある。

表-5 同一の決定規則が無い決定表

事例	条件属性				決定属性
	a	b	c	d	e
1	1	1	1	2	1
2	1	2	2	2	1
3	2	1	2	2	1
4	3	1	3	2	1
5	1	1	3	2	2
6	1	2	1	2	2
7	2	1	1	2	2
8	2	2	1	2	2
9	3	2	1	2	2

これは、表-2 の場合の条件属性集合を  $P$ 、決定属性集合を  $Q$  とすれば、

$$P = \{a, b, c, d\}, Q = \{e\}$$

となり、「条件属性  $\{a, b, c, d\}$  は決定属性  $\{e\}$  に従属するか否か。」すなわち「従属性  $P \Rightarrow Q$  が成立するか否か。」を決定する問題に帰着される。

例えば表-2 の事例 11 は、条件属性が  $P = \{1,1,1,1\}$  で、決定属性  $Q = \{2\}$  である。一方、事例 12~20 は、条件属性が同じ  $P = \{1,1,1,1\}$  であるが、決定属性  $Q = \{3\}$  と異なり、表-2 の決定表には矛盾がある。これを解消するためには、矛盾している決定規則を表-2 から全て排除するか、もしくは、表-2 を表-3 と表-4 に分離しなければならない。この場合、表-3 が矛盾の無い決定表であり、以降のステップでは表-3 の簡約化を行う。

[STEP-2] 同一規則のない決定アルゴリズムの作成

1つの決定アルゴリズム中に、同一の(全ての属性値が等しい)決定規則が複数存在しても意味が無いことから、これらをまとめて1つの決定規則とする。

表-3 の場合、事例 9 と事例 10 の決定規則は同一であることから、ここでは事例 9 の決定規則を残し、事例 10 を排除する。その結果、得られた決定表が、表-5 である。

[STEP-3] 条件属性集合の簡約化

条件属性集合の簡約化は、「従属性  $P \Rightarrow Q$  に対し、条件属性の集合  $P$  は不要な条件属性を含んでいるか否か。」すなわち「全体として決定表の識別能力を損なわずに、いずれかの条件属性を排除できるか否か。」という問題に帰着される。

条件属性集合の簡約化を行うためには、任意の条件属性集合を一旦決定表から排除し、その決定表すなわち決定アルゴリズムに矛盾が無いかを調べる必要がある。もし、矛盾が無ければ排除した条件属性は不要、反対に矛盾していれば不可欠とそれぞれ判断される。

具体的には、表-5 から最初に条件属性  $b, c, d$  を取り除

表-6 条件属性  $b, c, d$  を削除した決定表

事例	条件属性		決定属性
	$a$		$e$
1	1		1
2	1		1
3	2		1
4	3		1
5	1		2
6	1		2
7	2		2
8	2		2
9	3		2

表-7 条件属性  $c, d$  を削除した決定表

事例	条件属性		決定属性
	$a$	$b$	$e$
1	1	1	1
2	1	2	1
3	2	1	1
4	3	1	1
5	1	1	2
6	1	2	2
7	2	1	2
8	2	2	2
9	3	2	2

表-8 不要な条件属性を排除した決定表

事例	条件属性			決定属性
	$a$	$b$	$c$	$e$
1	1	1	1	1
2	1	2	2	1
3	2	1	2	1
4	3	1	3	1
5	1	1	3	2
6	1	2	1	2
7	2	1	1	2
8	2	2	1	2
9	3	2	1	2

いた決定表が表-6 となる。例えば、この決定表において、事例1と事例2、事例5と事例6の決定規則が同一規則となる。しかし、事例3と事例7,8の決定規則は矛盾しており、さらに、事例4と事例9の決定規則もそれぞれ互いに矛盾している。したがって、条件属性  $b, c, d$  には決定に必要なとされる属性が含まれていることとなり、排除できない属性である。

次に、表-5 から条件属性  $c, d$  を取り除いた決定表が表-7 となる。この場合、事例1と事例5、事例2と事例6、事例3と事例7の決定規則がそれぞれ互いに矛盾している。したがって、上記と同様、条件属性  $c, d$  は不可欠な属性あり、排除できない。以上の作業を、条件属性

全ての組み合わせ集合  $P'(\in P)$  について行う。

この結果、条件属性  $d$  を取り除いた場合、すなわち  $P' = \{a, b, c\}$  のとき、 $P'$  が  $(P, Q)$  における  $P$  の縮約となる。この場合の決定表が表-8 であり、矛盾する決定規則は一つも存在しない。つまり、表-8 は矛盾のない決定アルゴリズムとなる。

#### [STEP-4] 決定規則の縮約化

決定規則の縮約化は、STEP-3で縮約化された属性集合への射影による決定表において、「不要な条件属性値を取り除くことができるか否か。」という問題に帰着する。つまり、決定規則における不要な属性値の排除を行うためには、各々の条件属性の決定属性に関する核を計算する必要がある。言い換えると「条件属性のどの値が決定属性の値を識別するために不可欠であるか。」を決定することである。

核の値は、決定属性に対して不可欠な値の全ての集合であるから、ある値が不要で排除できるか否か、すなわち、その値が核に属するか否かを確認するためには、一旦その値を決定表から取り除き、同一行の残りの値により、その行の決定属性をただ一つに決定可能か否かを調べなければならぬ。もし、ただ一つに決定不可能な場合は、取り除いた値が核に属することとなる。

具体例として、表-8 の事例1の決定規則の核値を求めてみることにする。この決定規則において、条件属性  $a, b, c$  をそれぞれ取り除いた決定規則は、以下のようになる。

$$b=1, c=1 \rightarrow e=1$$

$$a=1, c=1 \rightarrow e=1$$

$$a=1, b=1 \rightarrow e=1$$

これらは、表-8 の他の決定規則とそれぞれ矛盾した決定規則であり、事例1の決定規則においては、 $a, b, c$  は排除不可能であり、いずれの条件属性値も核値となる。

次に、事例2の場合、この決定規則より、条件属性  $a, b$  の各値を取り除いた決定規則は、

$$b=2, c=2 \rightarrow e=1$$

$$a=1, c=2 \rightarrow e=1$$

であり、いずれも他の決定規則との矛盾が無く、 $a, b$  は排除可能なため、核値ではない。これに対し、条件属性  $c$  の値を取り除いた決定規則

$$a=1, b=2 \rightarrow e=1$$

は、矛盾した決定規則となり、事例2の決定規則において条件属性  $c$  の値が核値である。

このような一連の作業を、全ての決定規則における全ての条件属性値について行った結果、得られた各決定規則の核値を表-9 に示す。

表-9 各決定規則の核値

事例	条件属性			決定属性
	a	b	c	e
1	1	1	1	1
2	/	/	2	1
3	/	/	2	1
4	3	/	/	1
5	1	/	3	2
6	/	2	1	2
7	2	/	1	2
8	/	/	/	2
9	/	/	/	2

表-10 不要な属性値を排除した決定表

事例	条件属性			決定属性
	a	b	c	e
1	1	1	1	1
2	-	-	2	1
3	-	-	2	1
4	3	1	-	1
4'	3	-	3	1
5	1	-	3	2
6	-	2	1	2
7	2	-	1	2
8	2	2	-	2
8'	2	-	1	2
8''	-	2	1	2
9	3	2	-	2
9'	3	-	1	2
9''	-	2	1	2

表中の“-”は任意の値

次に、核値が決定された後、任意の決定規則において、核値に核値以外の条件属性値を加えた決定規則を考え、その規則が独立で、かつ矛盾が無いならば、その規則が決定規則の縮約となる。

表-8の事例1の場合、表-9より条件属性値の全てが核値となるため、

$$a=1, b=1, c=1 \rightarrow e=1$$

は独立かつ矛盾は無く、この規則そのものが事例1の決定規則の縮約となる。

一方、事例2の場合、核値である条件属性cの値に条件属性a, bの値を加えた決定規則が考えられる。aの値を加えた決定規則は、

$$a=1, c=2 \rightarrow e=1$$

となり、この規則からaの値を取り除いたとき、その規則には矛盾が無いため、独立でない。また、bの値を加えた決定規則は、

$$b=2, c=2 \rightarrow e=1$$

となり、この規則からbの値を取り除いたときも、その規則は矛盾が無いため、独立でない。ゆえに、この場合は条件属性値が核値のみで構成される

表-11 極小決定アルゴリズム

事例	条件属性			決定属性
	a	b	c	e
1	1	1	1	1
2,3	-	-	2	1
4	3	1	-	1
4'	3	-	3	1
5	1	-	3	2
6,8'',9''	-	2	1	2
7,8'	2	-	1	2
8	2	2	-	2
9	3	2	-	2
9'	3	-	1	2

表中の“-”は任意の値

$$c=2 \rightarrow e=1$$

が、事例2の決定規則の縮約となる。

次に、事例4の場合、核値である条件属性aの値に条件属性b, cの値を加えた決定規則が考えられる。bの値を加えた決定規則は、

$$a=3, b=1 \rightarrow e=1$$

となる。いま、この規則からbの値を取り除いたとき、その規則には矛盾が生じることから、bの値は不可欠な値となり、bの値を加えた決定規則は独立と言える。また、bの値を加えた決定規則には矛盾が生じない。続いて、cの値を加えた決定規則は、

$$a=3, c=3 \rightarrow e=1$$

となる。いま、この規則からcの値を取り除いたとき、その規則には矛盾が生じることから、cの値は不可欠な値となり、cの値を加えた決定規則は独立と言える。また、cの値を加えた決定規則には矛盾が生じない。ここで、上記2つの規則は独立かつ矛盾が無い規則であることから、この2つの規則が事例4の決定規則の縮約となる。

また事例8の場合、核値がないため、決定規則の組としてa, b, c, ab, ac, bcを考える必要がある。この決定規則の組のうち、独立かつ矛盾が無い規則が事例8の決定規則の縮約となる。表-10は、表-8における各決定規則の縮約を示したものである。

最後に、極小決定アルゴリズムを求めるため、表-10の複数存在する属性値が全て等しい同一の決定規則を1つの規則としてまとめ、表-11に示す極小な決定表が得られる。なお、図-2には、STEP4の流れを示している。

以上が、決定表の簡約化によるルール獲得すなわち極小決定アルゴリズム作成の手順である。本研究では、このラフ集合の概念による決定表簡約化手法のプログラムを用いて、伸縮継手の損傷に関する実データを対象としてルール型知識の獲得を試みた。

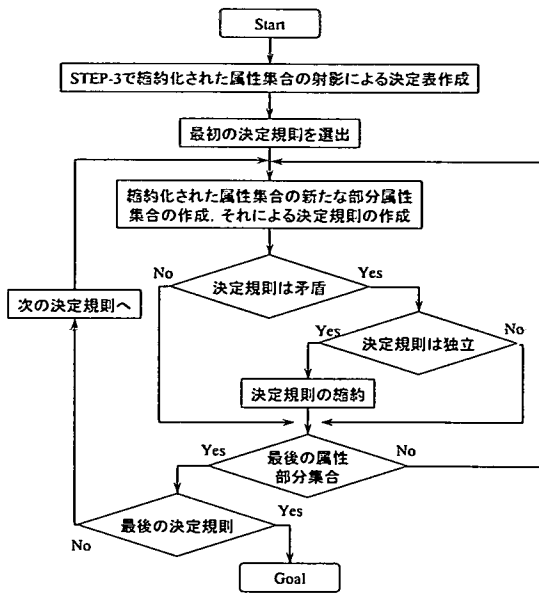


図-2 STEP-4のフロー

### 3. 獲得された知識の評価方法

本研究では、3つの評価法により、獲得されたルールを検証した。以下には、それぞれの評価法について記述する。

#### (1) 獲得された知識の識別率による評価法

本研究では、獲得した知識を評価する手法として、識別率による評価方法を提案する。ここで、識別率とは、獲得されたルール型知識が、ルール型知識を獲得する前のデータベースのデータに対して、どの程度の識別が可能であるかを示す指標である。以下には、ラフ集合の概念を利用した決定表簡約化手法に適用する「解析データ」と獲得された知識である「ルール」について定義した後に、識別率の式を示す。

解析データ：

「解析データ」とは、ラフ集合の概念を利用した決定表簡約化手法を適用する前のデータ類であり、矛盾および重複する決定規則を含むデータである。

ルール：

「ルール」とは、「解析データ」に対してラフ集合の概念を利用した決定表簡約化手法の[STEP-1]から[STEP-4]を適用し、獲得されたデータ類である。

解析データの識別率： $D_{analysis}$

獲得された「ルール」を用いた「解析データ」の識別可能な程度を、本研究では次式に示す識別率を用いて、

評価することとした。

$$D_{analysis} = \frac{\text{「ルール」によって識別できた「解析データ数」}}{\text{「解析データ」の全数}} \times 100 \quad (1)$$

「解析データ」の識別率  $D_{analysis}$  が 100%に近いほど、

「ルール」によって「解析データ」を正確に識別でき、「解析データ」をより反映したルール型知識が獲得できているといえる。

#### (2) アンケートによる評価法

ルール型知識の獲得結果を検証するため、専門技術者へのアンケートを実施した。アンケートの目的は、第一に獲得された知識に対する信憑性の評価、第二にデータマイニングという観点から、新たな知見としての知識獲得に対する正否の評価である。そこで、これら进行评估するために、本研究では妥当性、新規性、有用性、意外性の4つの評価指標を提案し、それぞれの定義を以下に示す。

妥当性 (Validity)：

専門知識との整合性。つまり、自分の経験や常識に合致しているかの指標。

新規性 (Freshness Degree)：

専門知識に存在しない知識。つまり、一度も経験の無いパターンかどうかの指標。

有用性 (Usefulness Degree)：

専門知識に対する有用性。つまり、専門分野における利用・応用に対する可能性の指標。

意外性 (Unexpected Degree)：

専門知識に対して部分的に反すること。つまり、専門知識と異なる度合いの指標である。

なお、アンケートにおける具体的な評価法は、獲得された知識に対して、専門技術者がそれぞれの評価指標について「ある」、「なし」の評価を行う。

#### (3) 獲得された知識の正答率による評価法

最後に正答率による評価方法を提案する。ここで定義する正答率とは、獲得されたルール型知識における決定部分と専門技術者の意見がどれだけ合致しているかという指標である。正答率による評価方法は、獲得されたルールの決定部分(決定属性の値)を「空白」にして、専門技術者が自らの考えで、その決定部分を書き込む手法である。その結果、獲得されたルールの決定部分と専門技術者が書き込んだ決定部分が一致していれば、常識的なルールと判断でき、一致していなければ、新たな知見のルールであると判断できる。

表-12 対象データテーブルとデータ項目数

データテーブル名		データ項目数
資産データ	道路構造	20
	上部工	36
	下部工	56
	伸縮継手	13
補修データ	伸縮継手補修	13
点検データ	伸縮継手点検	14

#### 4. 実橋データへの適用

##### (1) 伸縮継手の損傷に関するルール型知識獲得目的

本研究において橋梁の伸縮継手の損傷を対象とした理由は、橋梁構造物の種々の部位の中で、比較的損傷が多く、そのため補修工事の維持管理費用に占める割合が大きいことや<sup>10)</sup>、過去における解析対象となるデータが豊富なことから選定した。次に、ルール型知識を獲得する目的としては、第1に、伸縮継手の代表的な損傷<sup>17)</sup>である「異常音」、「漏水」、「錆」を対象としてルールを獲得し、それぞれの損傷に影響を及ぼしている重要な因子を特定する。第2に、データマイニングという知識発見の観点から、伸縮継手の損傷に関して専門技術者が有していない新たな知識を獲得することである。

##### (2) 対象データとその特徴

本研究において適用の対象としたデータは、橋梁の伸縮継手に関連するデータであり、阪神高速道路の橋脚上に設置されている伸縮継手に関するデータである。これらのデータは、阪神高速道路の効率的な維持管理業務の支援を目的として開発された、データベース<sup>10)</sup>に蓄積されているデータを用いた。データの特徴は、以下のとおりである。

###### a) 対象と下データテーブル

適用した伸縮継手に関連するデータは、上記データベースに蓄積されている30種類以上に及ぶデータテーブルの内、伸縮継手に関連する、道路構造資産、上部工資産、下部工資産、伸縮継手資産、伸縮継手点検、および伸縮継手補修の計6個のデータテーブルを選定した。

###### b) 各データテーブルの特徴

資産に関するデータテーブルは、構造物のしゅん工時に、対象とする構造物の維持管理に必要な事項を記録し、蓄積したデータであり、阪神高速道路の全資産のデータが蓄積されている。一方、補修、または、点検に関するデータテーブルも同様に、補修時、または点検時にそれぞれの必要事項を記録し、蓄積したデータであるが、これらのデータは過去、全てのデータが蓄積されていない

項目もある。ここで、表-12には、対象としたこれら各テーブルのデータ項目数を示す。

##### c) 612 橋脚分の伸縮継手データ

伸縮継手に関するデータテーブルの内、阪神高速道路のある路線の612橋脚上に設置されている伸縮継手を対象のデータとした。

##### d) 点検間隔は不規則

伸縮継手の点検は、1970年、1983年、1984年、1987年、1989年、1992年、1994年、1995年の計8回のデータであり、一定間隔での点検は行われておらず、点検間隔は不規則である<sup>18)</sup>。

##### e) 時系列データベース

④において述べたとおり、点検は1970年から1995年までの計8回行われており、その結果が蓄積された時系列データベースである。

##### (3) 適用データの前処理

前節で紹介したデータベースに蓄積されている伸縮継手に関連するデータの状態のままでは、データマイニングを行うことは困難である。そこで、これらのデータを「解析データ」に変換する必要がある、その処理の手法について以下に説明する。

###### a) 伸縮継手データ項目の選択

解析に用いる6個の伸縮継手に関連するデータテーブルを選定したが、これらの各テーブルは表-12に示すとおり詳細なデータ項目を有しており、これらの総計は94項目に及ぶ。しかしながら、これらの項目には、伸縮継手の損傷や構造には影響を与えない項目、例えば設置場所の住所や管理番号なども含まれる。そこで、コンピュータの処理能力も加味し、これらの項目の内、伸縮継手の設計条件となる項目、過去の事例より損傷に影響を及ぼすと思われる項目など、計11項目をデータマイニングに用いる項目として選定した。その結果、本解析の目的が「異常音、漏水、錆という伸縮継手の損傷に影響を及ぼす因子を特定すること」であるから、決定属性はそれぞれの「損傷判定結果」となり、条件属性は、「橋齢(しゅん工年)」、「縦断勾配」、「橋長」、「連続径間数」、「設計伸縮量」、「上下方向区分」、「平面線形」、「桁種別」、「伸縮継手種別」、「床版端部補強」、「伸縮継手供用年数」となる。

###### b) データ洗浄

データ洗浄とは、選定したデータにはいくつかの欠損値を含んでいることから、その削除、または修正を行うことである。この欠損値による解析精度の低下を回避する目的で、本研究では、基本的にデータの欠損値を含むレコードを、データベース上から削除した。また、誤入力されたと思われるデータについては、正しいデータを確認できる資料より、全て正しいデータ

表-13 各条件属性のカテゴリー区分

項目(条件属性)	0	1	2	3	4
A01 橋齢(しゅん工年)*	1970年	1971年	1972年	1973年	1976年
A02 縦断勾配	-5~2%	-1.9~-1.0%	-0.9~-0.1%	0%	0.1~0.9%
A03 橋長	10m未満	20m未満	30m未満	40m未満	50m未満
A04 連続径間数	0(単梁桁)	2	3	4	5
A05 設計伸縮量	20mm未満	40mm未満	60mm未満	80mm未満	100mm未満
A06 上下方向区分	上り	下り	2方向		
A07 平面線形	緩和曲線	R500m未満	R1,000m未満	R1,000m以上	
A08 桁種別	鋼桁	PC桁	RC桁		
A09 伸縮継手種別	ゴム継手	鋼製継手	突合せ継手		
A10 床版端部補強	無し	鋼板接着	ブラケット補強	併用補強	その他のタイプ
A11 伸縮継手供用年数	0年以上 2年未満	2年以上 4年未満	4年以上 6年未満	6年以上 8年未満	8年以上 10年未満
損傷判定結果**	Aランク	Bランク	Cランク		

項目(条件属性)	5	6	7	8	9
A01 橋齢(しゅん工年)*	1996年				
A02 縦断勾配	1.0~1.9%	2.0~2.9%	3.0~3.9%	4.0~5.9%	6.0~9.0%
A03 橋長	60m未満	80m未満	100m未満	130m未満	
A04 連続径間数	6				
A05 設計伸縮量	120mm未満				
A06 上下方向区分					
A07 平面線形					
A08 桁種別					
A09 伸縮継手種別					
A10 床版端部補強					
A11 伸縮継手供用年数	10年以上 12年未満	12年以上 14年未満	14年以上 16年未満	16年以上 20年未満	20年以上 26年未満
損傷判定結果**					

\* : しゅん工年は対象としたデータから、上記の5年のみ \*\* : ランク分けの内容は表-14に示す

表-14 損傷判定結果のランクの内容<sup>18)</sup>

損傷判定結果ランク	評価内容
A	損傷が著しく、早急に補修する必要がある場合
B	損傷があり、状況に応じて補修する必要がある場合
C	損傷が軽微である場合

表-15 入力データと事例数

データ種別	異常音	漏水	錆
決定属性・損傷判定結果の事例数	1,464	2,587	2,937

それら項目のカテゴリー区分を決定した。カテゴリーの区分の方法によっては、データマイニングの結果に大きく影響を及ぼすと考えられることから、対象データの数値範囲を調査し、例えば設計伸縮量については伸縮継手の形式に差異が生じない範囲などを考慮し、それぞれのカテゴリー区分を決定した。なお、カテゴリーデータに変換した項目は、橋齢(しゅん工年)、縦断勾配、橋長、設計伸縮量、伸縮継手供用年数の計5項目であり、表-13に適用した全項目のカテゴリー区分を示す。

e) 「解析データ」のコード化

「解析データ」のカテゴリー区分を決定した後、データのコード化を行った。本研究では、コード化を終えたデータを「入力データ」と呼ぶこととする。

f) 「入力データ」を用いたルール型知識の獲得

「入力データ」の作成後、このデータにラフ集合の概念

に修正した。

c) 「解析データ」の作成

目的とする「解析データ」を作成するために、表-12に示す6個のテーブルを結合し、時系列データを形成した。その結果、対象とした阪神高速道路の堺線における612橋脚分のデータから、8954件のレコードが生成された。

d) カテゴリー区分の決定

対象としたデータは、数値データとカテゴリカルデータが混在した状態となっている。そこで、データ形式を統一するために、選定したデータ11項目の内、数値データについてはカテゴリカルデータへ変換することとし、



表-16 縮約化された条件属性集合 (異常音)

属性数	[STEP-3]において縮約化された属性集合
7	A02,A03,A05,A07,A09,A10,A11
8	A01,A02,A03,A05,A07,A09,A10,A11
	A02,A03,A04,A05,A07,A09,A10,A11
	A02,A03,A05,A06,A07,A09,A10,A11
	A02,A03,A05,A07,A08,A09,A10,A11
9	A01,A02,A03,A04,A05,A07,A09,A10,A11
	A01,A02,A03,A05,A06,A07,A09,A10,A11
	A01,A02,A03,A05,A07,A08,A09,A10,A11
	A02,A03,A04,A05,A06,A07,A09,A10,A11
	A02,A03,A04,A05,A07,A08,A09,A10,A11
	A02,A03,A05,A06,A07,A08,A09,A10,A11
10	A01,A02,A03,A04,A05,A06,A07,A09,A10,A11
	A01,A02,A03,A04,A05,A07,A08,A09,A10,A11
	A01,A02,A03,A05,A06,A07,A08,A09,A10,A11
	A02,A03,A04,A05,A06,A07,A08,A09,A10,A11
11	A01,A02,A03,A04,A05,A06,A07,A08,A09,A10,A11

表-17 縮約化された条件属性集合 (漏水)

属性数	[STEP-3]において縮約化された属性集合
9	A02,A03,A04,A05,A07,A08,A09,A10,A11
10	A01,A02,A03,A04,A05,A07,A08,A09,A10,A11
	A02,A03,A04,A05,A06,A07,A08,A09,A10,A11
11	A01,A02,A03,A04,A05,A06,A07,A08,A09,A10,A11

表-18 縮約化された属性集合 (錆)

属性数	[STEP-3]において縮約化された属性集合
8	A01,A02,A03,A05,A07,A09,A10,A11
	A02,A03,A05,A06,A07,A09,A10,A11
9	A01,A02,A03,A04,A05,A07,A09,A10,A11
	A01,A02,A03,A05,A06,A07,A09,A10,A11
	A01,A02,A03,A05,A07,A08,A09,A10,A11
	A02,A03,A04,A05,A06,A07,A09,A10,A11
	A02,A03,A05,A06,A07,A08,A09,A10,A11
10	A01,A02,A03,A04,A05,A06,A07,A09,A10,A11
	A01,A02,A03,A04,A05,A07,A08,A09,A10,A11
	A01,A02,A03,A05,A06,A07,A08,A09,A10,A11
	A02,A03,A04,A05,A06,A07,A08,A09,A10,A11
11	A01,A02,A03,A04,A05,A06,A07,A08,A09,A10,A11

を用いた決定表簡約化手法 (以下、「本手法」と略記する) を適用した。「解析データ」には数多くの損傷データが含まれるが、本研究では、「異常音データ」、「漏水データ」、「錆データ」の3つ損傷別にデータセットを作成し、本手法を適用した。表-15には、各損傷の事例数を示す。

#### (4) 獲得された知識の考察

本節では、本手法によって獲得された知識、また、獲得された知識に対する専門技術者による評価結果について述べる。

表-19 各データの条件属性の出現頻度

条件属性		異常音	漏水	錆
A01	橋齢(しゅん工年)	8	2	8
A02	縦断勾配	16	4	12
A03	橋長	16	4	12
A04	連続径間数	8	4	6
A05	設計伸縮量	16	4	12
A06	上下方向区分	8	2	8
A07	平面線形	16	4	12
A08	桁種別	8	4	6
A09	伸縮継手種別	16	4	12
A10	床版端部補強	16	4	12
A11	伸縮継手供用年数	16	4	12

#### a) 縮約化された条件属性集合の評価

表-16~18には、本手法の[STEP-3]において不要な条件属性を削除して絞り込まれた条件属性集合の部分集合、すなわち、「異常音」、「漏水」、「錆」の各データにおける縮約化された条件属性集合を示す。なお、「異常音」は、条件属性集合が最も少ない7属性から11属性までを、「漏水」は、条件属性集合が9属性から11属性までを、「錆」は、条件属性集合が8属性から11属性までの組み合わせをそれぞれ示した。また、各データに対して、本手法の[STEP-3]において縮約化された条件属性集合は、以下のとおりである。

「異常音」：7属性

{縦断勾配, 橋長, 設計伸縮量, 平面線形, 伸縮継手種別, 床版端部補強, 伸縮継手供用年数}

「漏水」：9属性

{縦断勾配, 橋長, 連続径間数, 設計伸縮量, 平面線形, 桁種別, 伸縮継手種別, 床版端部補強, 伸縮継手供用年数}

「錆」：8属性

{橋齢(しゅん工年), 縦断勾配, 橋長, 設計伸縮量, 平面線形, 伸縮継手種別, 床版端部補強, 伸縮継手供用年数}

本研究では、条件属性の数が少ない部分集合ほど、実用面においては少ない要因で損傷の予測が可能であるなど、効率的であると考え、「異常音」、「漏水」、「錆」の各データにおける最も少ない属性数の属性集合に[STEP-4]を行った。また、条件属性数の最も少ない組み合わせから最大11属性の組み合わせまでを提示し、その出現回数についての考察を加えた。表-19には、それぞれのデータにおける表-16~18に示した条件属性集合全体における各属性の出現回数をそれぞれ計算した結果を示す。各データにより、出現回数の多い条件属性には多少のばらつきはあるものの、各データ出現回数が最も多い共通の条件属性(表-19中の網掛け部分)を挙げ

表-20 各損傷の事例数, ルール型知識の事例数,  
識別された事例数, 識別率の関係

データ	事例数 (a)	ルール型知識 の事例数	識別された 事例数(b)	識別率 (b/a)
異常音	1,464	426	1,271	87%
漏水	2,587	964	2,272	88%
錆	2,937	500	2,836	97%

ると、「橋長」、「設計伸縮量」、「平面線形」、「伸縮継手種別」、「床版端部補強」、「伸縮継手供用年数」が得られた。したがって、これら、6 属性が伸縮継手の損傷である「異常音」、「漏水」、「錆」のいずれにも大きな影響を及ぼしている重要な要因であると考えられる。これらの共通の条件属性に加え、各損傷に対してはそれぞれの出現回数の多い他の条件属性が、個別の損傷に影響を与える要因であると言える。

これらの要因は、過去分析された補修周期と損傷要因との事例<sup>19)</sup>と一部で同様の結果が得られているが、従前の解析事例は単一の要因との関係のみしか分析されておらず、本研究のように複数の要因の関係を明らかにできたことは、今後、これらの要因分析、あるいは効率的な予防保全を確立する上で、貴重な情報となり得る。

#### b) 識別率を考慮した評価

本手法によって獲得されたルール型知識が、実データである「解析データ」を精度良く反映しているかを検証するため、識別率による評価を行った。ここでは、その結果であるルール型知識の「解析データ」に対する識別率について考察する。

表-20 には、伸縮継手データにおける「異常音」、「漏水」、「錆」の各データの事例数、および、データマイニングによって獲得されたルール型知識の事例数、識別された事例数、および、識別率を示す。なお、表-20 における識別率は以下のように求めた。

「異常音」に関するルール型知識によって、「異常音」データの事例 1464 件のうち 1271 件が識別できていることから、

$$D_{\text{abnormal}} = \frac{1271}{1464} \times 100 = 87 (\%) \quad (2)$$

となる。以下同様に「漏水」、「錆」はそれぞれ

$$D_{\text{leak\_water}} = \frac{2272}{2587} \times 100 = 88 (\%) \quad (3)$$

$$D_{\text{rust}} = \frac{2836}{2937} \times 100 = 97 (\%) \quad (4)$$

となる。

上記より、「異常音」、「漏水」、「錆」に関するルール型知識を用いて識別された識別率を求めた結果、いずれの損傷の識別率も高いことから、元のデータを反映したルール型知識が獲得されていると考えられる。以下に、それぞれの考察を述べる。

「異常音」に関するルール型知識による識別率を求めた結果、87%であり、3つの損傷データの中では最も低い識別率となったが、残りの13%は、本手法における[STEP-1]および[STEP-2]において、矛盾する決定規則を取り除いたことから識別できなかった。しかし、90%近くが識別できていることから、ルール型知識の獲得は正確に行われたと言える。「漏水」に関するルール型知識の識別率についても88%であり、「異常音」の考察とほぼ同様の傾向といえる。最後に、「錆」に関するルール型知識による識別率は97%であったことから、識別率としては、3つのデータの中で最も高く、また、100%近く識別できていることから、ルール型知識の獲得はかなり正確に行われた。

以上の結果から、各損傷に関するルール型知識は、「解析データ」の属性数よりも少ない属性数で識別できること、また「解析データ」の事例数よりも少ないルールの数で識別できることから、効率的なルール型知識が獲得されたと考えられる。

#### c) アンケートによる評価<sup>20)</sup>

本研究は、獲得されたルール型知識を検証するため、専門技術者による評価を導入した。評価方法は、3.(2)で述べた4つの指標で評価するアンケートであり、その詳細、および、アンケート結果から得られた各損傷の傾向について述べる。なお、実施したアンケートは、獲得されたルール型知識の事例に対し、「異常音」は156問、「漏水」は111問、「錆」は116問をそれぞれ選別し、橋梁維持管理に関わる経験年数が3年~20年の専門技術者、延べ26人に対して実施した。表-21~23には、各損傷別のアンケート結果から得られたルール型知識の各評価指標に対する傾向を示す。

これらの結果を見ると、多くの専門技術者が共有している伸縮継手の損傷に対する要因は、伸縮継手の供用年数がいずれの指標においても大きい。つまり、この年数が短いにもかかわらず損傷判定結果が悪い場合は、意外性があるとともに新規性が高くなり、場合によっては有用性も高い。これに対し、伸縮継手の供用年数が長くて損傷判定結果が悪い場合は妥当性が高くなっている。しかし、新規性、および、意外性の高いルール型知識は、従来の専門技術者が有していた知識のみでは判断しにくい要因が発見されている可能性がある。例えば新規性の評価で見られた橋長や平面線形の

表-21 「異常音」に関するルール型知識の各評価指標に対する傾向

評価指標	各指標の評価が高いルール型知識の傾向
妥当性	<ul style="list-style-type: none"> <li>条件属性「伸縮継手供用年数」が2年未満で「損傷判定結果」がCランクであるルール群。また、同じく8年以上10年未満でAランク、またはBランクであるルール群。つまり通常、伸縮継手の補修は約10年以上経過後であるため、「伸縮継手供用年数」が長いほど、「損傷判定結果」が悪くなるのは常識的な考えである。</li> </ul>
新規性	<ul style="list-style-type: none"> <li>妥当性の高いルール群とは逆に、条件属性「伸縮継手供用年数」が2年未満や4年未満と短いにもかかわらず「損傷判定結果」がAやBであるルール群。</li> <li>条件属性「橋長」が50m以上60m未満で、「平面線形」が緩和曲線もしくはR=1,000m以上の直線に近い線形区間で「損傷判定結果」がAやBであるルール群。これは、橋長が長く、伸縮量（遊間）が大きな伸縮継手となり、かつ線形が直線に近いため車両の走行速度が速く、伸縮継手に対する衝撃が大きくなることが理解できるが、このような、線形条件と伸縮量を併せて判断する知識は少ない。</li> </ul>
有用性	<ul style="list-style-type: none"> <li>新規性とと同じく、条件属性「伸縮継手供用年数」が2年未満や4年未満と短く、かつ「伸縮継手種別」が耐久性が比較的高い、鋼製継手やゴム継手にもかかわらず、「損傷判定結果」がAやBであるルール群。「損傷判定結果」を「伸縮継手供用年数」のみから判断するのではなく、他の影響因子を分析することにより、早期劣化に与える影響を把握できる可能性がある。</li> </ul>
意外性	<ul style="list-style-type: none"> <li>有用性とと同じく、条件属性「伸縮継手供用年数」が2年未満や4年未満と短く、「伸縮継手種別」が耐久性が比較的高い、鋼製継手やゴム継手にもかかわらず、「損傷判定結果」がAやBであるルール群。早期劣化の傾向が見られるルール群はいずれも平均的に意外性が高い。</li> </ul>

表-22 「漏水」に関するルール型知識の各評価指標に対する傾向

評価指標	各指標の評価が高いルール型知識の傾向
妥当性	<ul style="list-style-type: none"> <li>条件属性「伸縮継手供用年数」がいずれの年数であっても「損傷判定結果」がBランクであるルール群。これは、「漏水」は「伸縮継手供用年数」にかかわらず、比較的多く発生する損傷である。</li> <li>条件属性「伸縮継手種別」が鋼製継手で、かつ排水型のタイプでは、「伸縮継手供用年数」がたとえ2年未満であっても「損傷判定結果」がAランクの場合、このタイプにおいては排水樋の損傷が多いため、常識的な考えであり、意外性は少ない。</li> </ul>
新規性	<ul style="list-style-type: none"> <li>「異常音」の場合とほぼ同様に、条件属性「伸縮継手供用年数」が2年以上4年未満、「伸縮継手種別」が鋼製継手で「損傷判定結果」がBランクであるルール群。妥当性の場合とは異なり、比較的耐久性のある鋼製継手では新規性が高い。</li> </ul>
有用性	<ul style="list-style-type: none"> <li>妥当性と同じく、条件属性「伸縮継手供用年数」がいずれの年数であっても「損傷判定結果」がBランクであるルール群。これは、「漏水」は「伸縮継手供用年数」が短くても比較的多く発生する損傷であり、その発生のパターンは把握しやすい。</li> <li>条件属性「伸縮継手種別」が突合わせ継手で、「伸縮継手供用年数」が20年以上26年未満で「損傷判定結果」がAランクであるルール群。これは、かなり耐久性が高いと判断できる結果・条件であり、「漏水」に対する耐久性把握に有益である。</li> </ul>
意外性	<ul style="list-style-type: none"> <li>妥当性とほぼ同様であるが、「伸縮継手種別」が耐久性が高い鋼製継手にもかかわらず、「伸縮継手供用年数」が2年未満で「損傷判定結果」がAランクであるルール群。妥当性もあるが、専門知識には部分的には反することから、意外性が高い。</li> </ul>

要因がこれにあたり、このことは合理的な維持管理計画を立案する上で、貴重な情報の発見ができていものと考えられる。

#### d) 獲得された知識の正答率による評価結果

ここでは、3.(3)に記述された、「本手法によって獲得されたルール型知識」と「専門技術者が有する知識」がどの程度一致するかを調査した知識の正答率について考

察する。

本研究では、伸縮継手の代表的損傷である「異常音」、「漏水」、「錆」に関するルール型知識を獲得した。ここで、獲得したルール型知識の決定属性である「損傷判定結果」を空白にしたアンケートを、専門技術者に対して行った後、獲得したルール型知識の「損傷判定結果」と専門技術者の知識で解答した「損傷判

表-23 「錆」に関するルール型知識の各評価指標に対する傾向

評価指標	各指標の評価が高いルール型知識の傾向
当性	<ul style="list-style-type: none"> <li>条件属性「伸縮継手供用年数」が6年以上で「損傷判定結果」がAランク、Bランク、Cランクであるルール群。これは「錆」のみであれば、6年以上経過すれば、いずれのレベルの損傷が発生しても不思議でない状況である。</li> </ul>
新規性	<ul style="list-style-type: none"> <li>条件属性「伸縮継手種別」がゴム継手、もしくは突合せ継手で、「伸縮継手供用年数」が4年未満で「損傷判定結果」がBランクやCランクであるルール群。錆の損傷が少ないゴム継手、もしくは突合せ継手においては4年未満での錆は想定し難く、新たな要因が考えられる。</li> </ul>
有用性	<ul style="list-style-type: none"> <li>条件属性「伸縮継手供用年数」が6年以上で、「損傷判定結果」がCランクであるルール群。これは、妥当性とは反対に6年以上経過しているにもかかわらず、軽微な損傷であり、耐久性向上の要因が把握できる可能性がある。</li> <li>条件属性「伸縮継手種別」が突合せ継手で「伸縮継手供用年数」が8年以上10年未満で「損傷判定結果」がAランクであるルール群。これは、錆の損傷が少ないタイプにもかかわらず損傷が激しいことから、耐久性の把握に有益である。</li> </ul>
意外性	<ul style="list-style-type: none"> <li>条件属性「伸縮継手種別」がゴム継手にもかかわらず、「伸縮継手供用年数」が2年未満で「損傷判定結果」がCランクであるルール群。新規性にもあるとおり、2年程度での損傷は、専門知識には反する。</li> <li>条件属性「伸縮継手種別」がいずれのタイプにかかわらず、「伸縮継手供用年数」が12年以上にかかわらず「損傷判定結果」がCランクであるルール群。10年以上経過しているにもかかわらず、損傷が軽微であるため、意外性が高い。</li> </ul>

表-24 正答率によるアンケート結果

損傷種別	設問数	正答率アンケート結果内訳	
異常音	99問	専門技術者より評価が悪いルール型知識の数	92
		専門技術者と評価が一致したルール型知識の数	6
		専門技術者より評価が良いルール型知識の数	1
		正答率	6%
漏水	111問	専門技術者より評価が悪いルール型知識の数	49
		専門技術者と評価が一致したルール型知識の数	36
		専門技術者より評価が良いルール型知識の数	26
		正答率	32%
錆	116問	専門技術者より評価が悪いルール型知識の数	27
		専門技術者と評価が一致したルール型知識の数	38
		専門技術者より評価が良いルール型知識の数	51
		正答率	33%

注) 「専門技術者よりも評価が悪いルール型知識の数」とは、獲得されたルール型知識の「損傷判定結果」がBランクに対し、専門技術者が解答した「損傷判定結果」がCランクと低い場合を指し、逆に「専門技術者よりも評価が良いルール型知識の数」とは、獲得されたルール型知識の「損傷判定結果」がBランクに対し、専門技術者が解答した「損傷判定結果」がAランクと高い場合を指す。

定結果」の一致する割合を調べた。表-24には、各損傷のアンケート内容、およびその結果を示す。

表-24からは、それぞれの損傷に関して、アンケートの結果から得られた正答率が6%、32%、33%である。このことは、専門技術者の知識を反映したルール型知識が獲得されているとはいえない。しかし、逆に正答率が100%もしくは、100%に近い結果であった場合、専門技

術者のもつ知識だけが獲得されていることを意味し、データベース内に潜在的に存在し、専門技術者が気づいていないような未知のルール型知識が獲得されていないことになる。したがって、データマイニングの観点からは、正答率が低ければ、専門技術者が有していない新たな知見であるルール型知識が獲得されていると考察できる。

詳細に結果を評価すると、獲得したルール型知識と專

門技術者の解答が一致したルール型知識は、常識的なルール型知識と考えられ、従来、専門技術者が有していた知識と同等の知識と考えられる。一方、獲得したルール型知識の評価が専門技術者の解答よりも悪いルール型知識は、専門技術者が考えるよりも、別の要因が加わることで、劣化が予想以上に進行していると考えられる。このことは、予防保全の観点から非常に重要な知見が獲得されていると考えられ、得られたルール型知識の条件属性である損傷要因を今後十分に分析する必要がある。最後に、獲得したルール型知識の評価が専門技術者の解答よりも良いルール型知識は、逆に専門技術者が考えたほど、劣化が進行していないと考えられる。

以上の評価から、劣化進行に関して新たな知見を得られたことは、今後の維持管理に非常に有益な情報、劣化に影響を与える損傷要因等を得たこととなり、点検の重点項目の設定や補修計画の立案など応用範囲は様々である。

## 5. 結論

本研究は、橋梁の維持管理に関する新たな知識を獲得することを目的として、既存データベースより伸縮継手の損傷に関するルール型知識の獲得を試みたものである。さらに本研究では、獲得されたルール型知識を評価するための検証方法を提案し、獲得された知識の検証を行った。以下には、本研究で得られた成果をまとめる。

- ① ラフ集合の概念を利用した決定表簡約化手法を用いて、ルール型知識を獲得するプログラムを開発した。本プログラムを実際に供用されている伸縮継手データに適用することによって、「異常音」、「漏水」、「錆」という3種類の損傷に関する簡約化された決定表が作成でき、効率的なルール型知識が獲得できた。
- ② 獲得されたルール型知識によって、「解析データ」である「伸縮継手データ」を識別した結果、高い識別率が得られたことから、「伸縮継手データ」を反映したルール型知識が獲得できたといえる。また、「伸縮継手データ」の属性数よりも少ない属性数、および、「伸縮継手データ」の事例数よりも少ないルール型知識の数で識別できていることから、ルール型知識の獲得は成功したといえる。
- ③ 新たな知識のルール型知識が獲得されているかを検証するため、妥当性、新規性、有用性、意外性を指標としたアンケートによる評価を提案し、実際に専門技術者に対してアンケートを行った。その結果、本研究で提案した評価指標は、どの程度従来の知識

と一致するか、また、どの程度新たな知識が獲得できているかを評価することが可能となり、検証するための指標として有効であることが示された。

- ④ 専門技術者へのアンケート結果から、妥当性、新規性、有用性、意外性のあるルール型知識の傾向を把握することができた。その際、早期劣化のパターンや、劣化速度の遅いパターンにおける影響因子が把握できたことから、本研究で提案した評価指標は、ルール型知識に対する評価を把握するための指標としても有効であることが示された。損傷に対する影響因子が把握できたことにより、今後の合理的な維持管理を確立する上で、有益な情報として活用できるものと考えられる。
- ⑤ 正答率によるアンケートを実施し、獲得されたルール型知識を評価した結果、専門技術者の評価との一致、不一致により、新たな知見が獲得できているかを検証するための方法論として有効であることが明らかとなった。

今後は、データのさらなる充実をはかり、ルール獲得の精度を向上させるとともに、新たに得られた知見の有効活用について検討を行いたい。

謝辞：本研究をまとめるにあたり、種々のご援助を頂いた山口大学 中村秀明助教授、および山口大学大学院生 河上和志氏に感謝の意を表します。

## 参考文献

- 1) Pieter Adriaans・Dolf Zantinge 著、山本英子・梅村恭司 訳：データマイニング、共立出版、1998.6.
- 2) 河野浩之：知識発見とデータマイニング、日本ファジィ学会誌、Vol.9, No.6, pp.851-860, 1997.12.
- 3) 福田剛志、森本康彦、森下真一、徳山 豪：データマイニングの最新動向、情報処理、Vol.37, No.7, pp.597-603, 1996.
- 4) 河野浩之：データベースからの知識発見の現状と動向、人工知能学会誌、Vol.12, No.4, pp.497-504, 1997.
- 5) 田中 淳：隠れた情報を見いだすデータマイニングへの挑戦、日経コンピュータ、pp.91-98, 1995.1.23.
- 6) ハーブ・エデルスタイン：データマイニングで眠れる情報を掘り起こせ、日経コンピュータ、pp.187-190, 1996.3.18.
- 7) 吉田琢也：情報の脈脈を掘り当てるデータマイニングいよいよ本番、日経コンピュータ pp.70-73, 1996.9.16.
- 8) 森 健次郎：急速に普及するデータ・マイニング安価で易しいツールが起爆剤に、日経コンピュータ、pp.188-189, 1998.1.19.
- 9) 沼尾雅之、清水周一：流通業におけるデータマイニング、

人工知能学会誌, Vol.12, No.4, pp.528-535, 1997.

- 10) 横森 貴, 小林 聡: ラフ集合と意思決定, 数理科学 No.375, サイエンス社, pp.76-83, 1994.9.
- 11) 古田 均, 広兼道幸, 田中成典, 三雲是宏: 橋梁の損傷要因診断事例からのラフ集合を用いたルール型知識の獲得方法, 構造工学論文集, Vol.44A, pp.521-528, 1998.3.
- 12) 古田 均, 木村壽夫, 広兼道幸, 田中成典, 三雲是宏: 橋梁の損傷要因診断事例からのラフ集合を用いたルール型知識の獲得および共有方法に関する研究, 構造工学論文集, Vol.45A, pp.533-541, 1999.3.
- 13) 広兼道幸: 斜面の崩壊危険度診断における観測データ分析手法に関する研究, 広島大学博士論文, 1996.3.
- 14) Pawlak, Z.: *Rough Sets, International Journal of Computer and Information Science*, Vol.11, No.5, pp.341-356, 1982.
- 15) 中村 昭: ラフ集合—その基本概念と知識情報, 数理科学 No.373, サイエンス社, pp.78-83, 1994.7.
- 16) 加賀山泰一, 丹波寛夫, 黒崎剛史, 柴田洋: 阪神高速道路における保全情報管理システムの構築, 橋梁と基礎, Vol.35, pp.24-30, 2001.12.
- 17) 加賀山泰一, 丹波寛夫, 柴田洋, 多田成宏: 阪神高速道路におけるライフサイクル評価(その2) —伸縮装置の評価—, 第56回年次学術講演会, 2001.10.
- 18) 阪神高速道路公団: 道路構造物の点検標準(土木構造物編), 1996.5.
- 19) 加賀山泰一, 広瀬鉄夫, 丹波寛夫: 阪神高速道路におけるライフサイクル評価の一手法, 阪神高速道路公団技報, Vol.20, 2002(投稿中)
- 20) 鈴木英之進: 共通データからの仮説駆動型例外ルール発見, 人工知能学会誌, 15巻, 5号, pp.782-789, 2000.9.

(2002.7.22受付)

## RULE-TYPE KNOWLEDGE ACQUISITION FROM BRIDGE EXPANSION JOINT DATABASE BY USING ROUGH SET BASED DATA MINING METHOD

Taiichi KAGAYAMA, Kei KAWAMURA, Ayaho MIYAMOTO and Shinya TANAKA

In recently, some useful database systems have been developed for civil infrastructures maintenance management. In this study, the acquisition of the rule type knowledge concerning the bridge expansion joint deterioration is attempted by using the Knowledge Discovery and Data Mining (KDDM) technique from the databases. As the KDDM technique, the decision table abridgment, which used the concept of the rough sets, is applied to the extraction of the rule-type knowledge. Then it offers useful tools to discover patterns on the bridge expansion joint damages hidden in data stored in bridge-database. And also, several different domain experts have evaluated the results obtained by the KDDM technique.