

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

1.はじめに

エキスパートシステムの開発にあたり知識獲得は
システム構築上, 最も困難なプロセス

1.初期における専門知識の体系化

- ・多くが手作業であり多大な労力を要する
- ・経験的知識であるため知識が複雑

2. 知識の追加・修正の過程

- ・追加・修正されたルールの信頼性の評価
- ・ルール間の整合性を保つ

3.経験的知識であることによる知識の曖昧さ

- ・知識の真偽を定量的に評価することが困難

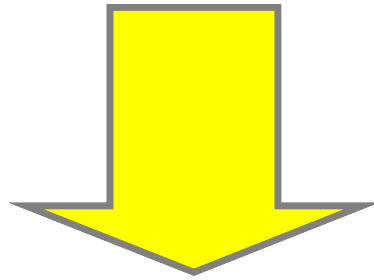
本研究では、

対象問題：鋼橋疲労損傷の補修方法選定

- 1.知識の共有，再利用を目的
知識洗練機能付推論システムが
対象ドメインによらず機能するかを検討
- 2.ルールベース構築支援を目的
平均圧縮情報量による
データベースからのルール抽出アルゴリズム
を提案し，その有効性を検討

知識洗練機能付推論システム

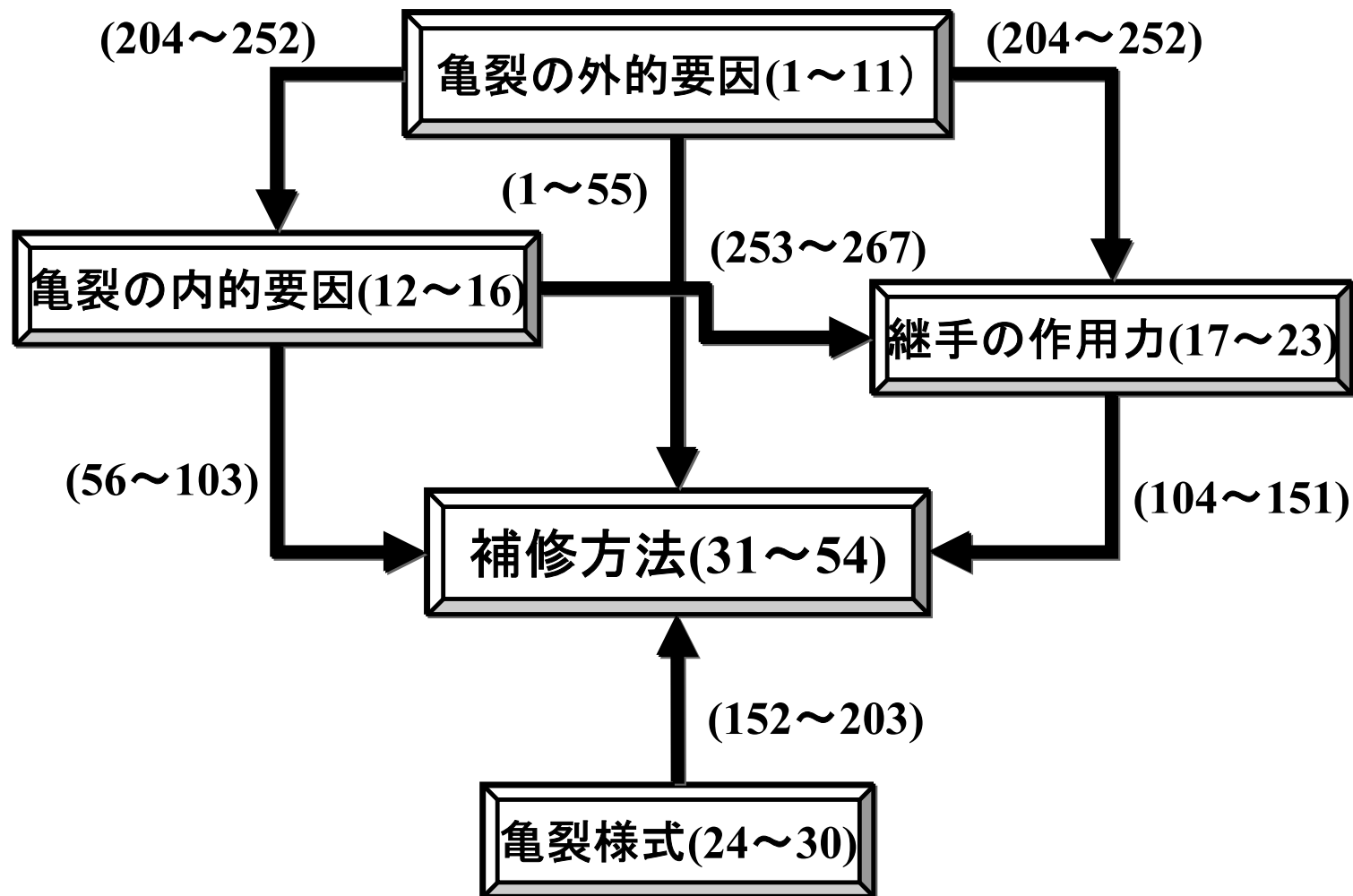
- min-max演算
- ニューラルネットワークで用いられる誤差逆伝播アルゴリズムを応用した推論方法
- 確信度の付与



- 通常, エキスパートシステムで行う仮説推論
- 推論システムが持つルールベースの洗練

適用したルールベース

田中らにより開発された鋼道路橋疲労損傷の補修方法選定システムのルールベース

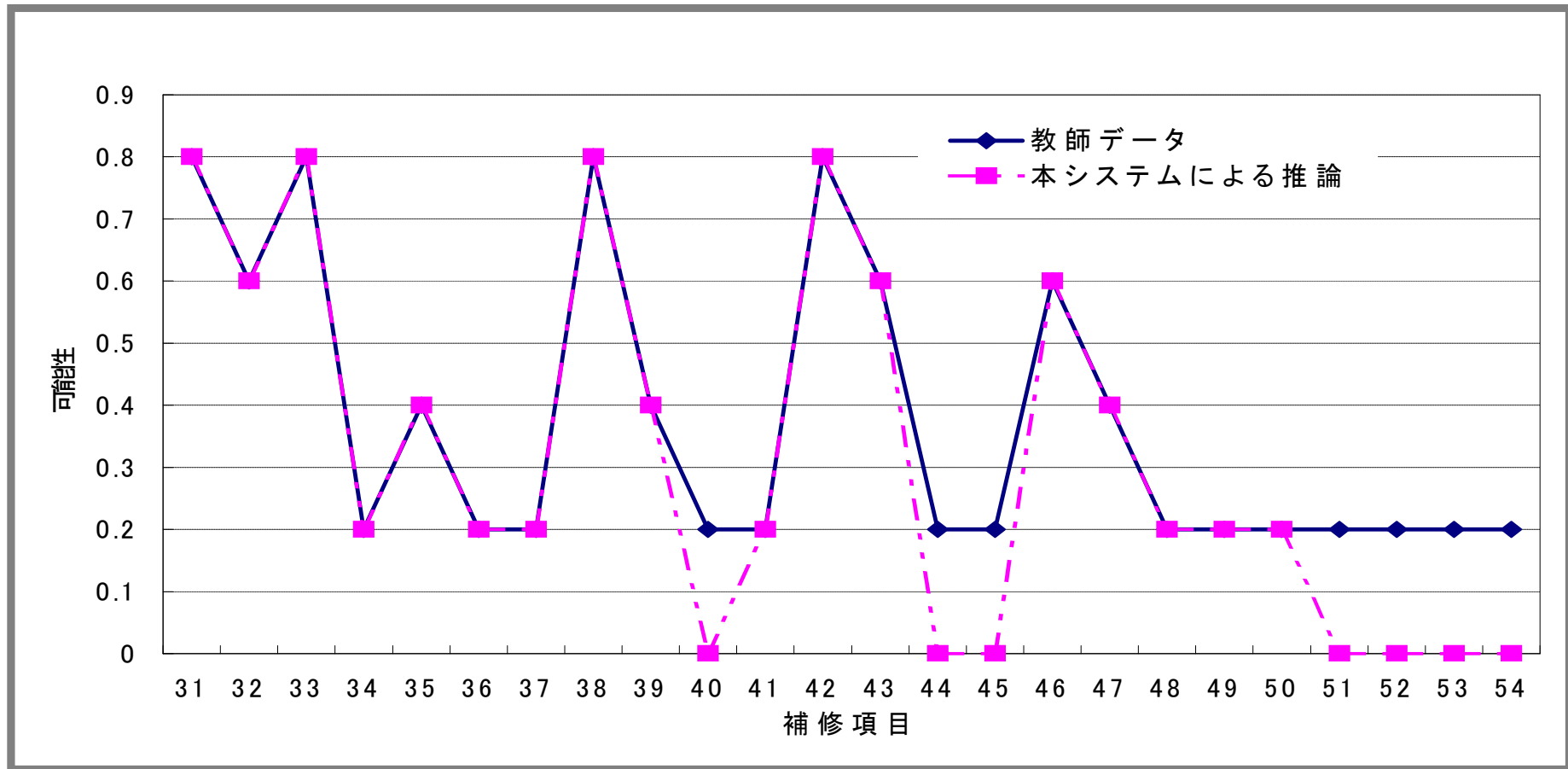


推論システムの適用結果

提示する事例

橋梁名	事例1	事例2
国名 架設年度 損傷発見年度	米国 1958年 12年	米国 1973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 溶接種類 継手種類 損傷要素	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 隅肉溶接 重ね継手 隅肉溶接	主桁の垂直補剛材取り付け部 主桁腹板 垂直補剛材 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 損傷状況	ピーニング 再溶融 高力ボルト 添接板	ストップホール グラインダ

補修方法選定結果(事例2)



- ・実際にとられた補修方法のストップホール、
グライндаを最も高い可能性としている
- ・推論全体の傾向が教師データにほぼ一致している

ルールベース構築支援を目的とした 平均圧縮情報量によるルール抽出

知識洗練機能付推論システムでは
ある程度の妥当性の確認された
既存の知識ベースの枠内での
知識洗練による知識獲得

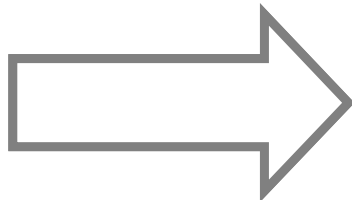
新たな知識を発見するための方法論として
情報理論の分野で用いられている
平均圧縮情報量の知識発見への適用性を検討

データベース中の要素x
についての情報量

$$-\log p(x)$$

$$-\log p(\bar{x})$$

ルール $y \rightarrow x$
を抽出

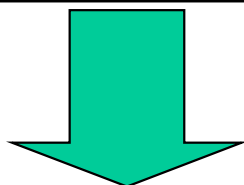


要素xについての
減少した情報量

$$-\log p(x | Y)$$

$$-\log p(\bar{x} | Y)$$

情報量の差をyがもたらす情報量

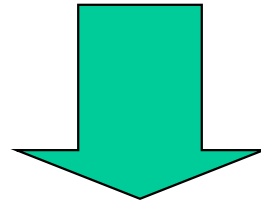


平均圧縮情報量: ACE

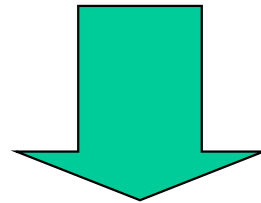
$$ACE(x, Y_\mu) = p(x, Y_\mu) \log \frac{p(x | Y_\mu)}{p(x)} + p(\bar{x}, Y_\mu) \log \frac{p(\bar{x} | Y_\mu)}{p(\bar{x})}$$

仮想事例によるデータベースの作成

鋼橋疲労損傷の補修方法選定システムのルールベース



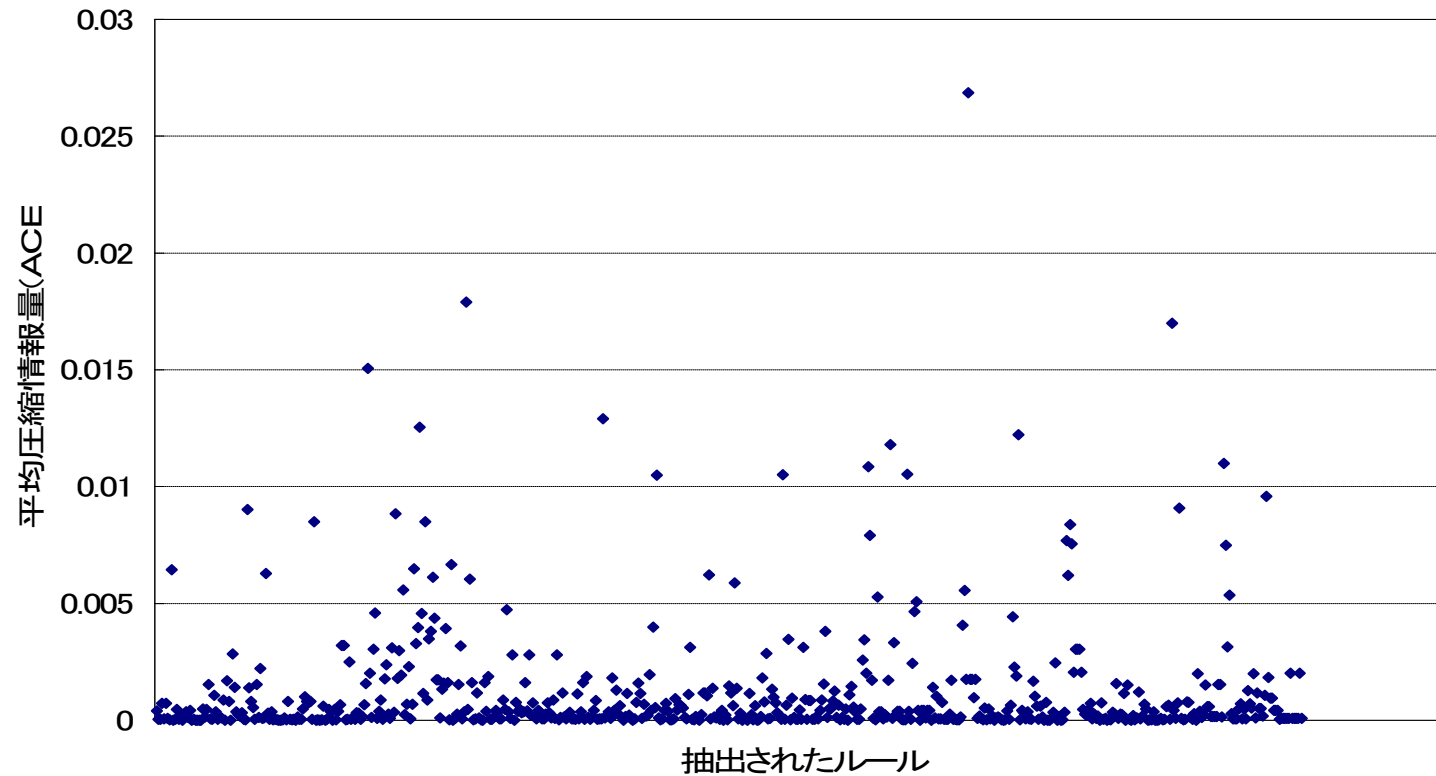
亀裂の内的要因, 外的要因, 継手の作用力, 亀裂様式
について起こり得る可能性のある全組み合わせ
2695通りを仮想事実



ルールベース洗練機能付推論システムより
補修方法を唯一つ選定して仮想事例

データベースからのルール抽出

平均圧縮情報量を用いたルール抽出アルゴリズムより
ルール抽出を試みた結果、620のルールが抽出された

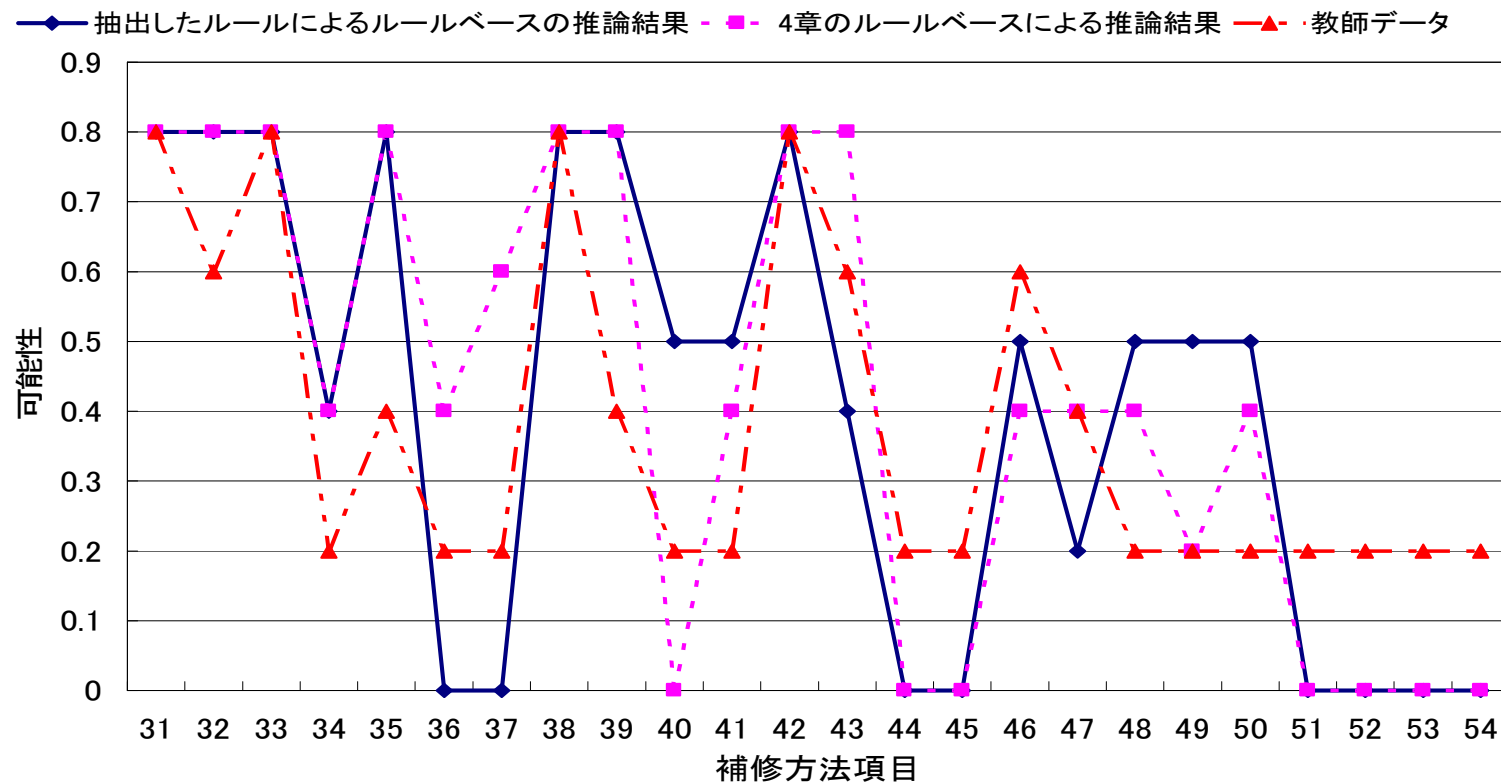


抽出されたルールの有用性の検討

抽出されたルールを基にルールベースを構築

- a) 全ルールを5分割し，上位40%の結合係数を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%毎に結合係数を0.8, 0.6, 0.4, 0.2とした場合
- c) 全ルールを10分割し，10%毎に結合係数を0.8から0.1までとした場合

b) の場合の補修方法選定結果 (事例2)



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、
十分な可能性の分類を行って推論結果を提示している

おわりに

1. 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
2. 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
3. 抽出したルールを用いて実際の事例を適用して
補修方法選定を試み、抽出したルールの有用性を検討