

帰納学習をとり入れた意思決定支援システムの設計\*

4H-1

大和田 勇人<sup>†</sup>, 大熊 喜之<sup>†</sup>, 杉本 勉<sup>†</sup>, 溝口 文雄<sup>†</sup>

<sup>†</sup>東京理科大学, <sup>‡</sup>NTT データ通信(株)

1 はじめに

戦略的情報システム (Strategic Information Systems, 以下 SIS) とは, 競争相手に対し差異化を図り優位に立つための情報システムで, 環境の変化に対応して増収, 増益を実現することがその役割である。

近年の市場競争の激化に伴い, 大企業ばかりではなく中小企業にも SIS が次々に導入されている。中小企業を対象にした SIS 開発の場合, それにかかるコストが重要である一方, 市販ツールの単なるチューニングでは同業他社との差異化を図ることが困難になるという問題がある。個々の SIS にオリジナリティがあり, なおかつその開発において高い生産性を実現するためには, コンサルタントと工程の標準化および支援システムの協調が有効に機能しなければならない。そこで NTT データでは, SIS 開発のための標準化手法として「SCAW デザインシリーズ」を作成し, 現在そのシステム化を進めている。ここではその標準化手法における開発の一工程である SFB (Structured Frame method for Business systems design) に焦点を当てる。これは, 事業戦略や業務改善の提示がされた後のシステム企画および設計を行うフェーズのことであり, 設計型エキスパートシステムと見ることもできる。

本研究では, SFB に対して帰納学習機構を導入することにより, 過去のシステム開発例からルールや関係などを導出する。学習手法には述語論理式を生成する帰納的論理プログラミング [1] に基づき, 学習ツールである GOLEM [2] を用いる。この方法により, SIS の生産性を向上させることが可能となる。

なお, 本システムは UNIX 上に構築されており, データの入力はスプレッドシート形式にて行う。

2 方法

2.1 入力データ

入力データは, SFB において会社のどの部門に SIS を導入すべきかの判断を行うために使われる。ここでは以下に示す3種類のデータを扱う。

\*A decision support system with inductive reasoning.

<sup>†</sup>Science University of Tokyo

<sup>‡</sup>NTT DATA Communications System Corp.

1. SFB の為の背景知識.

これは, 表 1, 2 に示すような SFB において対象となる会社の経営状態の調査結果である。

表 1: 各業務の経営状態

会社名	業務領域	経営状況
丸菱商事	営業管理	一部済
...	...	...

表 2: 営業課題の重要性

会社名	経営課題	重要性
丸菱商事	コスト削減	最重要
...	...	...

2. SFB の事例.

これは, 表 3 に示すような調査結果から導き出された各業務部門ごとの情報システムの導入に関する優先度である。

表 3: 情報システム導入の優先度

会社名	業務部門	優先度
丸菱商事	営業部門	最優先
丸菱商事	財務部門	優先
...	...	...

3. SFB と矛盾する事例.

3つめは前に述べた優先度のデータと矛盾するようなデータである。これは, 帰納学習において過剰な一般化を防ぐために必要なものである。

2.2 帰納学習

上で示したデータから, SFB のための一般規則を導出する。導出は第 1 階述語論理式の部分集合を一般化する帰納学習システム GOLEM[2] を用いる。

表 4: 矛盾する事例

会社名	業務部門	優先度
三胃	営業部門	優先
丸菱商事	財務部門	最優先
...	...	...

GOLEM は背景知識  $K$  のもとで正事例集合  $E^+$  の  $RLGG$  (Relative Least General Generalization) を求める。  $RLGG$  は  $E^+$  を導出する論理式の中で最も特殊なものであり, GOLEM は事例集合からボトムアップに一般規則を導出する。但し  $RLGG$  は負事例集合  $E^-$  を導出しない。

一般に “論理式  $A$  が  $C$  に関して  $B$  より一般的である” ことを以下のように示す。

$$C \models A \rightarrow B \text{ かつ } C \not\models B \rightarrow A$$

これより,  $RLGG$  は以下のように定式化される。

1.  $K \models (RLGG \rightarrow E^+)$
2.  $K \not\models (E^+ \rightarrow RLGG)$
3.  $K \models (C \rightarrow E^+)$   
なる無矛盾なすべての  $C$  において  
 $K \models (C \rightarrow RLGG)$
4.  $K \not\models (RLGG \rightarrow E^-)$

ここでは, 上で示したデータを GOLEM に対して以下のように与える。

- $K = \{SFB\text{の爲の背景知識} \cup SFB\text{の事例}\}$
- $E^+ = SFB\text{の事例集合}$
- $E^- = SFB\text{と矛盾する事例集合}$

### 3 実験

入力データは SFB の為の背景知識については 12 種類, 事例については 4 種類を 6 社分与え, 矛盾する事例については 10 例与えた。学習に要した時間は約 4 秒であった。以下に, 学習結果によって得られた規則とその意味を示す。

1. 次式は, 経営課題の顧客コミュニケーションが最重要ならば, 営業部門のシステム化が最優先であることを示している。

$$\forall x(\text{顧客コミュニケーション}(x, \text{最重要}) \rightarrow \text{営業部門}(x, \text{最優先})).$$

2. 次式は, 経営課題のコスト削減が最重要ならば, 営業部門の優先度が最優先であることを示している。

$$\forall x(\text{コスト削減}(x, \text{最重要}) \rightarrow \text{営業部門}(x, \text{最優先})).$$

3. 次式は, 経営課題のコスト削減と顧客コミュニケーションの重要性が等しい会社は, 財務部門と人事部門の優先度が等しいことを示している。

$$\forall x, y, z(\text{コスト削減}(x, z), \text{顧客コミュニケーション}(x, z), \text{人事部門}(x, y) \rightarrow \text{財務部門}(x, y)).$$

実際に, この学習結果を入力データと照らし合わせてみる。式 (3) について考えると, 表 2,3 を見ればわかるように丸菱商事に関してこの式が正しいことがわかる。

このように帰納学習を行なうことによって人間ではわかりづらい経営状態と情報システムの導入の優先度との関係が明らかになる。さらに異なった業務部門間の関係も把握することができる。

## 4 おわりに

本稿では, SIS 開発工程の一部に帰納学習を組み込んだ実験について述べた。

今後は, SIS に対する問題として, より複雑な入力データに適応できる方式を開発すること, 帰納学習の問題として学習結果の解釈の与え方や Prolog との結合による学習結果の正当性の証明などについて検討していく予定である。

### 謝辞

本研究を行うにあたり協力して頂いた NTT データ通信 (株) SIS 推進本部川崎洋輝担当課長および鈴木努担当課長代理に感謝いたします。

### 参考文献

- [1] S. Muggleton. Inductive logic programming. In *New Generation Computing*, 1991, Vol 8, pages 42-61, 1990.
- [2] S. Muggleton and C. Feng. Efficient induction of logic programs. In *Proc. First Conf. on Algorithmic Learning Theory*, pages 369-381, 1990.