

社会情報解析：

概念・関係・論理からのアプローチ

Social Information Analysis:
an approach through concept, relation and logic

長田 博泰

1. はじめに

本稿は2012年8月28日開催の「知の創成シンポジウム」で行った口頭発表を加筆・補綴したものである。当日は発表時間も限られ、目的や意図を十分に伝えられなかった。そもそもタイトルの「社会情報解析」という言葉が耳慣れない言葉であり、一般に必ずしも受け入れられていないのだから、狙いが伝わらないのは考えてみれば当然であろう。

本稿の構成は以下のとおりである。次節で「社会情報解析」のねらいとその方法のひとつである概念・関係・論理からのアプローチがいかなる関係にあるかを「ホーソン実験」に適用し考察する。ホーソン実験を取り上げる理由はつぎの二つにある。ひとつは、説明のための「人工的」例ではなく、社会学・経営学・心理学等の分野でよく知られる社会的具体例であること、いまひとつは、分析に用いられた統計的方法・ソシオグラム等と概念・関係・論理と比較・対比しやすいからである。3節以降は形式・関係・論理解析の具体的展開・説明であり、講演を整理・補綴したものである。

2. 社会情報解析—目的と方法

自然科学の現象は科学者の考察以前に解釈されてはいないし、現象を構成する物質、た

とえば原子や分子や電子にとって現象が何かの「意味」をもつということはない。しかし、社会科学の現象は、社会科学者の考察に先立ち、そこで生活している人々によって既に解釈されているのである。したがって、社会科学においては「純粋な事実」といったものは存在しない。

社会科学的現実がこのようなものだとするのなら、社会科学が経験科学として社会認識の客観性を保つためには、社会現象を把握する何らかの方法が必要となる。その有力な手段が社会調査であり、例えば質問紙調査に基づく研究では調査対象を多くし、大数の法則に基づいた分析が可能である。一方、個別的・典型的事例を重視する研究では、観察あるいは聞き取り調査等が用いられる。この場合、データ量も少なく、観察者自身の主観性も入りやすくなる。前者は計量的研究、後者は質的研究である。

人と人、人と社会のダイナミックな関係の中で情報のもたらすさまざまな働き・効果を研究対象とする社会情報学には、当然情報現象を分析し、法則化・予測等を行う方法を欠くことはできない。分析には従来の主たる分析方法である量的分析および質的分析が必要となる。しかし、少なくとも既成の方法には還元できない意味的・論理的分析が求められる。なぜなら、情報は送り手と受けての間で多様に変化する価値判断を含んでいるからで

ある。このような自覚の下で社会情報学の分析にはこのような視点が必要であると考え提唱するのが「社会情報解析」である。

このような価値判断を意味的・論理的に分析する新たな方法のひとつが概念・関係・論理にもとづくアプローチである。これは数理的方法ではあるが、統計的数量的方法ではない。しかも量的データにも質的データにも適用することができる。近年、情報端末等の普及に伴い膨大な電子データの中から新たな知見を見出すことや知識発見等に関心がもたれているが、この方法はその目的にも有効な方法である。

2.1 ホーソン実験とホーソン効果

[Homans, 進藤勝美]

1924年11月から1932年6月にかけてシカゴ郊外のホーソンにあるウェスタン・エレクトリック社のホーソン工場で行われた職場における作業効率や生産性を左右する要因について研究するために行われた実験が、いわゆるホーソン実験である。この研究から作業効率の向上には、温度や照明のような作業環境や休憩時間などの制度的条件よりも作業員間の非公式な人間関係がより重要な要素であることがはじめて発見された。

一方、この研究に対しては、実験対象の大部分が女性であった、作業者の〈自分たちには期待がかかっている〉という意識が研究者にとって都合のよい結果を生み出している可能性がある（これを「ホーソン効果」という）、労働組合がまったく考慮されていない、ことなどへの批判もあったが、従来の〈従業員はお金のためだけに働くものだ〉という単純な考え方が反省され、従業員の社会的欲求の充足を基調とする、いわゆる「人間関係論的管理」を生み出すことになる。

80年以上前のホーソン実験の統計分析およびその解釈をめぐるはいまも種々活発な議論が行われている。これは人間行動の評価

がいかに難しいものであるかを示すものであって、さらに人間行動の研究・評価方法の展開が期待されるというべきであろう。

2.2 ホーソン実験の分析と概念・関係・論理解析

ここで注目したいことは上のような結論を導出するために開発・適用された分析方法がほとんどそのまま概念・関係・論理解析へ翻訳できることである。言い換えれば、ホーソン実験の解析に用いられるソシオグラム、隣接行列、所属行列、二部グラフなどは概念・関係・論理解析の概念・関係の枠組みで一元的に扱うことが可能である。以下、これを具体的に示そう。

ホーソン実験関係の文献で分析対象となる基本データは、実験参加者基礎データ、友人・対立・助け合い・職務交換などの人間関係そして生産高で測られる作業効率などである。ここで分析対象とするデータを表1、表2に掲げる。

ホーソン実験の文献で取り上げられていない学歴、勤続年数および生産高と友人関係の関連をここでの方法、すなわち形式概念と関係概念を用いて分析する。そのために学歴・勤続年数・生産高を表3に示すようにカテゴリデータに変換する。

表3は対象と属性：学歴・勤続年数・生産

表1 配線作業観察室データ

工員		年齢	学歴	勤続(月)	生産高
検査工	I 1	23	中卒	36	
	I 3	40	専卒	84	
巻線工	W 1	22	小卒	38	724
	W 2	25	中卒	65	860
	W 3	26	小卒	29	823
	W 4	20	中卒	43	757
	W 5	24	中卒	32	804
	W 6	21	中卒	37	826
	W 7	22	小卒	38	651
	W 8	22	中卒	44	710
	W 9	21	小卒	34	416
熔接工	S 1	21	小卒	64	
	S 2	26	小卒	110	
	S 4	20	小卒	36	

表2 実験参加者の友人関係

友人関係	I 1	I 3	W 1	W 2	W 3	W 4	W 5	W 6	W 7	W 8	W 9	S 1	S 2	S 4
I 1					×									
I 3														
W 1					×	×						×		
W 2														
W 3	×		×			×						×		
W 4			×		×							×		
W 5														
W 6														
W 7										×	×	×		
W 8									×		×			×
W 9									×	×				×
S 1			×		×	×			×					
S 2														
S 4										×	×			

×印は友人関係を表す

高の関係である。一方、友人関係は対象間の関係である。表3の対象と属性の関係（概念関係—後述）を図示すると以下のようになる。

さて、図1は工具×属性関係であるが、これに工具間の友人関係を結び付けると属性と友人関係の関連がみえてくると予想される。この予想に沿って形式概念解析に関係概念を結びつけたものが関係概念解析である。複数の属性と関係概念を同時に解析すると図が複雑になるので、ここではそれぞれの属性と友人関係に関係概念解析を適用しよう。

1) 学歴と友人関係

図2をどのように解釈すればよいであろうか。図2のノード1, 2, 3はそれぞれ中学

校、専門学校、小学校卒を表す。ノード7は友人関係を表すノードであり、これが小・中・専卒と結合した結果できるノードが5と6である。この結果、友人関係の外延 {I 1, S 1, S 4, W 1, W 3, W 4, W 7, W 8, W 9} が小・中卒に分けられている。なお、この集合に入っていないI 3, S 2, W 2, W 5, W 6は友人関係にない。

2) 勤続年数と友人関係

勤続年数と友人関係も上述の学歴と同様の見方で説明することができる。一人を除いて全く同じ工具から構成されていることがわかる。

表3 対象×属性表（クロス表）

工具	学歴			勤続年数(月)			生産高		
	小卒	中卒	専卒	短い	中	長	低い	普通	高い
I 1		×		×					
I 3			×			×			
W 1	×			×				×	
W 2		×			×				×
W 3	×			×					×
W 4		×			×			×	
W 5		×		×					×
W 6		×		×					×
W 7	×			×			×		
W 8		×			×			×	
W 9	×			×			×		
S 1	×				×				
S 2	×					×			
S 4	×			×					

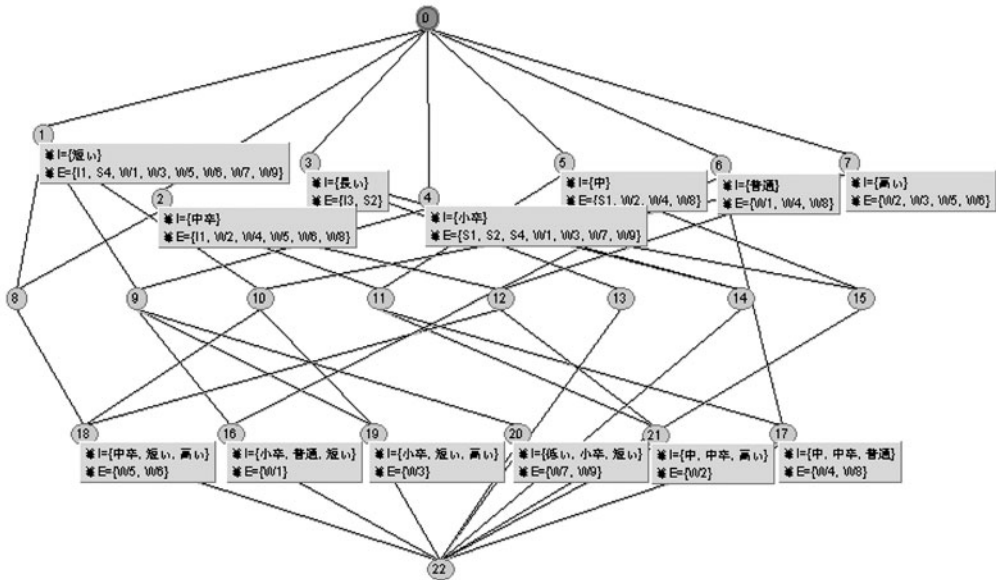


図1 表3の対象×属性の関連図(概念束)

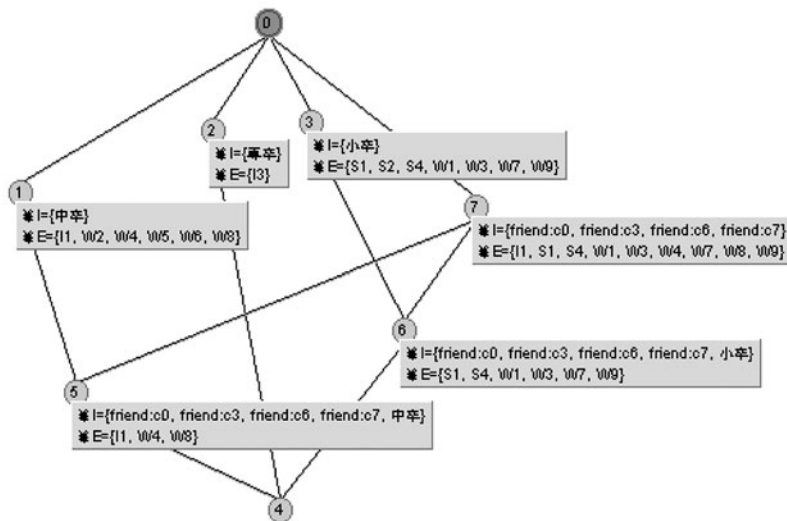


図2 学歴と友人関係概念図

3) 生産高と友人関係

ホーソン実験に参加した調査スタッフは生産高に影響をおよぼす要因や集団行動の他の特徴の指標として生産高を使うことに興味をもっていた。そして知能テストや器用テストのどちらの結果も平均時間当り生産高と相関しないことを見出した。実際、W9は、知能テストでは第一位に格付けられ、生産高では

最小であった。またW2は、知能テストでは最低位に格付けられ、生産高では最高であった。

その生産高と友人関係の関連を表しているのが図4である。図4の属性を見てわかるように生産高は学歴や勤続年数に比して極めて複雑であり、さらに生産高が高い工具のうちW2, W5, W6が友人関係をもたないので

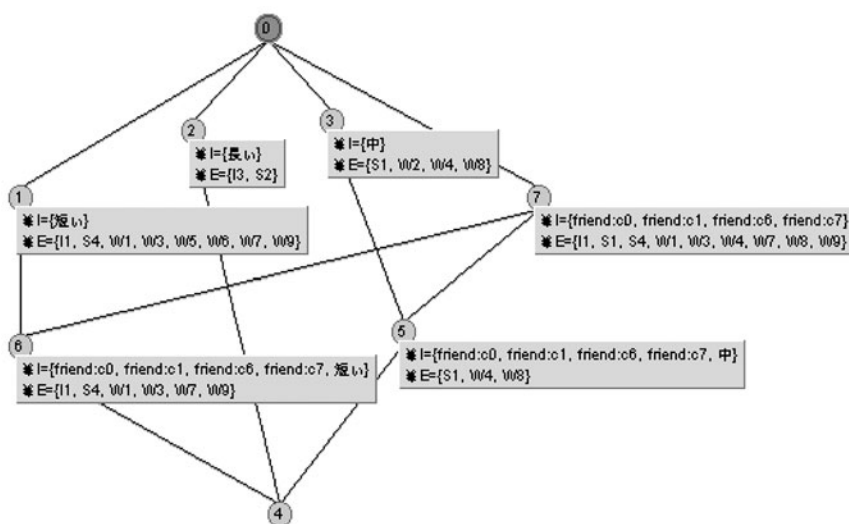


図3 勤続年数と友人関係概念図

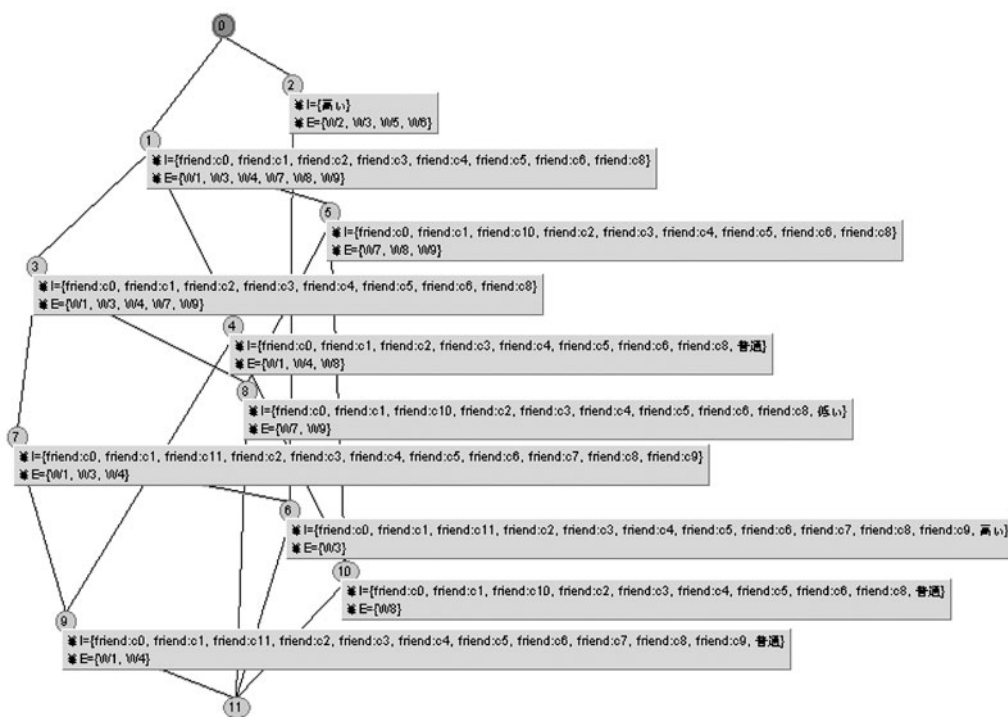


図4 生産高と友人関係概念図

概念束の上位にしか現れていない。したがって、生産高に友人関係が強い関連をもっているとは考えにくい。

学歴、勤続年数および生産高の各々と友人関係の関連を述べたが、学歴、勤続年数と友

人関係には親和性が認められるが、生産高に関して友人関係との関連が小さいと判断される。したがって、生産高の高い工具がかかわってくる工具間の別種の関連を調べる必要がある。事実、ホーソン実験ではW2, W5, W

6 が関わっている工具内の対立関係、ゲーム参加あるいは助け合いなどの関係も取り上げている。

しかし、ここはホーソン実験の再分析が目的ではなく、むしろ種々の特性と関係概念を結合する概念・関係・論理概念解析が有効な方法であることを理解、ないしは実感してもらうことが目的であるから、本題の形式・関係・論理概念解析の説明に移る。

3. 概念・関係・論理からのアプローチ

概念・関係・論理からのアプローチは、社会現象の多様性、関係の複雑さを捉えるため・統計的数量的方法によらず、
・データ間の特性に注目し、
・その内在的構造を明らかにすることを目的とする。まず対象と対象の有する属性（特性）から構成する形式概念について解説し、次いで対象と対象の「関係」と形式概念を結合する関係概念に拡張し、最後に対象の記述を属性から論理式による表現を可能にする論理概念解析について述べる。

3.1 形式概念解析 (Formal Concept Analysis—FCA)

万人が納得する概念の定義を与えることは至難である。ここでは 17 世紀のポール・ロワイヤル論理学に基づく概念の定義法を採用する。すなわち概念をその外延の集合と内包の集合の一致をもって定義する。ここで、外延とは「素数 = {2, 3, 5, ...}」のように集合とその元を枚挙することである。他方、内包とは「素数 $x = "x$ は 1 より大きい正の整数で、1 と自分自身以外に約数をもたない"」のように対象の共通性質を記述することである。

このように定義した概念を形式概念といい、つぎのような特徴を有する：

- ・概念の哲学的考察の数学的扱いである：対象を属性によって記述する
- ・対象を構造化し、分析する人間中心の方法

である：比較・順序に依拠した自然な方法である

- ・対象とその内在的構造、含意および依存関係を可視化する方法である：属性の包含関係だけから構造化する数理的手法である。

1) 形式文脈・形式概念・概念束 (がいねんそく)

形式概念解析の方法を太陽系惑星 (2006 年 8 月以前) の属性表によって説明しよう。形式概念解析はスライド 5 に示すような対象 O と属性 A からなる表—クロス表—からはじまる。行と列の交わったエントリに "x" が記されていれば、対象が列の先頭行で示される属性をもつことを意味し、対象と属性の二項関係を I と記す。三つ組み (O, A, I) を形式文脈という。

形式文脈から形式概念をつぎのように作成する (スライド 6)。対象の任意の集合をとる。例えば、 $X = \{\text{地球}\}$ とする。つぎに X に含まれる対象が共有する属性の集合をつくり、これを X' で表わそう。 $X = \{\text{地球}\}$ であれば、 $X' = \{\text{小, 近い, 衛星有}\}$ である。つぎに属性の集合 X' のすべての属性を有する対象の集合を見つける。これを X'' と記す。いまの場合、 $X'' = \{\text{地球, 火星}\}$ である。 X'' と X' の対を形式概念、対象の集合 X'' を外延、属性の集合 X' を内包という。 $(X'', X') = (\{\text{地球, 火星}\}, \{\text{小, 近い, 衛星有}\})$ が形式概念である。

形式文脈に含まれるすべての形式概念を見出すアルゴリズムがあり、最も基本的な方法はつぎのとおりである：

アルゴリズム：

文脈 (O, A, I) の概念のすべての外延を見つける。

- (1.1) 属性と外延という 2 つの列からなる表 (以下、 $[m|m']$ 表という) を作る。属性列の最初のセルを空にし、外延列の最初のセルに O を書く。

- (1.2) クロス表から属性の極大な外延を見つける(残っている属性の中で外延が最も大きいものを選ぶ), これを m' とする.
- (1.2.1) m' がまだ外延列に現れていないなら, $[m|m']$ 表に行 $[m|m']$ を付け加える. 集合 m' と外延列の既にあるすべての外延と共通集合を求める. この共通集合を外延列に(リストになれば)付

け加え, 属性列のセルを空にする.

- (1.2.2) m' がすでに外延列に現れているなら, m' がある行の属性セルに m を追加する.
- (1.3) クロス表から属性 m の列を削除する.
- (1.4) 最後の列が削除されたなら, 終了. さもないと, (1.2) に戻る.

太陽系惑星の属性記述

A: 属性の集合

	小	中	大	近い	遠い	衛星有	衛星無
水星	×			×			×
金星	×			×			×
地球	×			×		×	
火星	×			×			
木星			×		×	×	
土星			×		×	×	
天王星		×			×	×	
海王星		×			×	×	
冥王星	×				×	×	

対象O, 属性Aと関係Iの組 (O, A, I) を形式文脈という.

2012/8/28

知の創成シンポジウム

5

形式概念の例

	小	中	大	近い	遠い	衛星有	衛星無
水星	×			×			×
金星	×			×			×
地球	×			×		×	
火星	×			×			
木星			×		×	×	
土星			×		×	×	
天王星		×			×	×	
海王星		×			×	×	
冥王星	×				×	×	

1) 対象の任意の集合Xをとる, e.g. $X = \{\text{地球}\}$

2) $X' = \{\text{小, 近い, 衛星有}\}$

3) $(X')' = \{\text{小, 近い, 衛星有}\}' = \{\text{地球, 火星}\}$

4) $(X'', X') = (\{\text{地球, 火星}\}, \{\text{小, 近い, 衛星有}\})$

2012/8/28

知の創成シンポジウム

6

記法の導入

- Oの部分集合Xに対し,
Xの対象に共通な属性の集合をX'で表す.
 $X' =_{\text{def}} \{a \in A \mid \forall o \in X \text{ に対し } o/a\}$
- 同様に, Aの部分集合Yに対し,
Yのすべての属性をもつ対象の集合をY'で表す.
 $Y' =_{\text{def}} \{o \in O \mid \forall a \in Y \text{ に対し } o/a\}$

2012/8/28

知の創成シンポジウム

7

下位概念, 上位概念, 概念束

- (X_1, Y_1) と (X_2, Y_2) を文脈の概念とする. $X_1 \subseteq X_2$ (あるいは $Y_2 \subseteq Y_1$) なら, (X_1, Y_1) は (X_2, Y_2) の下位概念であるという. このとき (X_2, Y_2) は (X_1, Y_1) の上位概念であり, $(X_1, Y_1) \leq (X_2, Y_2)$ と書く.
- 関係 \leq を概念の階層順序(あるいは単に順序)という. この順序の概念集合の全体を $\mathcal{C}(O, A, I)$ で表し, 文脈 (O, A, I) の概念束という.

2012/8/28

知の創成シンポジウム

8

概念束

I

	小	中	大	近い	遠い	衛星有	衛星無
水星	×			×			×
金星	×			×			×
地球	×			×		×	
火星	×			×			
木星			×		×	×	
土星			×		×	×	
天王星		×			×	×	
海王星		×			×	×	
冥王星	×				×	×	

- 形式概念 $(X, Y) = (\{\text{水, 金, 地, 火, 冥}\}, \{\text{小}\})$ をとる
他の形式概念 $(X_1, Y_1) = (\{\text{水, 金, 地, 火}\}, \{\text{小, 近い}\})$
形式概念 (X_1, Y_1) は (X, Y) の下位概念, (X, Y) は (X_1, Y_1) の上位概念である.

2012/8/28

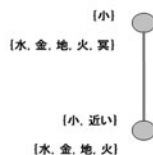
知の創成シンポジウム

9

概念束: Hasse図

II

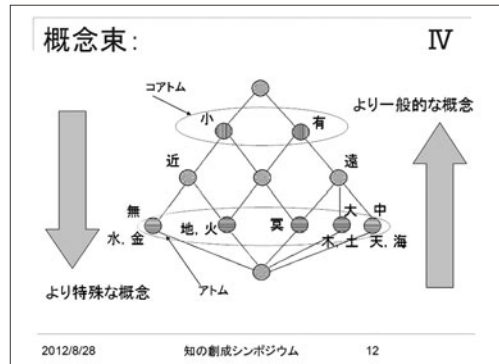
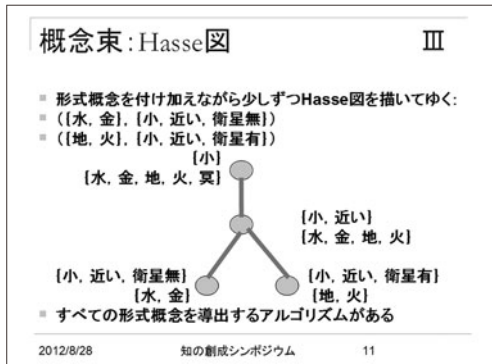
$(\{\text{水, 金, 地, 火}\}, \{\text{小, 近い}\})$ を $(\{\text{水, 金, 地, 火, 冥}\}, \{\text{小}\})$ の下に描き, 線で結ぶ



2012/8/28

知の創成シンポジウム

10



形式概念の間にスライド 8 に示すような順序関係を定義することができる。すなわち 2 つの形式概念 $(X1, Y1)$, $(X2, Y2)$ の順序 $(X1, Y1) \leq (X2, Y2)$ を外延間の包含関係, $X1 \subseteq X2$ と定義する (内包でいえば, $Y2 \subseteq Y1$ である)。この関係は概念の上位, 下位関係に相当するものであり, 形式概念でもこの用語を借用する。 $(X2, Y2)$ は $(X1, Y1)$ の上位概念, 逆に $(X1, Y1)$ は $(X2, Y2)$ の下位概念である。上位概念を図の上に描き, 下位概念の図の下に描き, 両者を線分で結ぶ。この描画をすべての形式概念の上位・下位概念に適用するとひとつの図ができる。これを Hasse 図という。こうして描いた太陽系惑星の Hasse 図がスライド 12 である。

2) 適用例—東京 23 区職種構成クラスタリング [長田]

量的データを質的データに変換し形式概念解析を適用する例として, スライド 18 の東京 23 区の職種区分別人数から 23 区のカテゴリをとりあげる。スライド 18 のデータを質的データに変換する手順は, 通常の統計的方法と変わらない。すなわち, スライド 18 の各区および合計の職種構成をそれぞれ各区人口合計, 合計人口で割って, 構成比を求める。さらにそれぞれの合計職種構成比で各区職種構成比を割り算し, 構成比の相対比 (= 特化係数) を計算する。この結果を特徴パターンで表現した結果をスライド 21 に示す。

この後の形式概念解析手順は上述のとおりであるので詳細は省略する。結果を Hasse 図に描いたものがスライド 22 である。

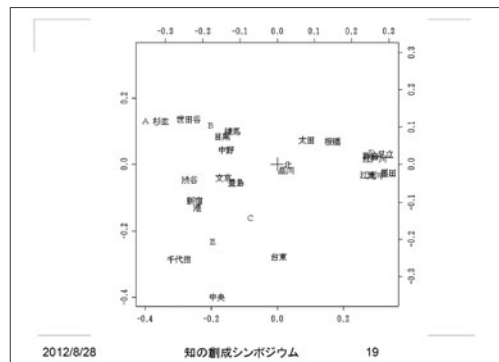
スライド 22 からつぎのことがわかる。すな

表5 東京23区(住民の職種構成)(1995年)

地域	A	B	C	D	E	合計
千代田	8	10	15	14	10	57
中央	7	13	24	22	14	80
港	20	27	24	40	23	134
新宿	33	46	39	65	37	220
文京	19	28	25	45	17	134
台東	14	24	43	66	25	172
江東	14	23	31	99	15	182
豊田	10	27	30	111	17	195
品川	24	45	37	99	27	232
目黒	23	36	23	52	16	150
大田	43	77	56	184	35	395
豊田	62	85	53	102	32	334
渋谷	22	34	25	42	22	145
中野	27	46	34	66	20	193
杉並	48	68	46	64	22	248
豊島	23	44	37	68	25	197
北	21	51	40	99	24	235
荒川	9	21	26	82	12	150
板橋	21	45	34	118	19	237
練馬	30	44	31	70	18	193
足立	15	38	37	143	18	251
葛飾	15	36	33	124	18	226
江戸川	13	31	30	112	18	204
合計	521	859	773	1887	452	4562

A: 管理職職種 B: 事務職職種 C: 販売職職種
D: 工場労働者 E: サービス職職種

2012/8/28 知の創成シンポジウム 18



あるいは&)で結合された論理式の含意(ならば, \rightarrow あるいは \Rightarrow)に限られる:

例: 衛星無 \Rightarrow 小, 近い \equiv 衛星無 \Rightarrow (小 \wedge 近い)

- ・否定および選言(または, \vee)を表現することができない

これらの問題点を解決するために形式概念解析を拡張する. 対象間の関係を含む形式概念解析が関係概念解析であり, 属性記述と論理表現の拡張するための形式概念解析方法が論理概念解析である. これら二つの方法を順に説明する.

3.2 関係概念解析 (Relational Concept Analysis—RCA)

ホーソン実験の例からわかるように各個人固有の属性を有するだけでなく, 各人が何らかの繋がりをもちそれがまた各個人の属性に反映する. このような現実を考慮するなら, 対象と属性の対だけを考察対象とする形式概念は余りに狭いといわざるを得ない. 各人の関係, 一般には対象間の関係を考慮し, その関係を取り込んだ形式概念解析が関係概念解析である. それは以下の特徴を有する:

- ・概念間の関係を導入し, これを形式概念と結合したものである.
- ・関係概念解析は主につぎの三つの処理からなる:

—実体関連モデル (Entity-relationship Model)を基礎とする関係モデルである.

—対象間の関係を関係属性として表現する概念的尺度を有する.

—逐次的処理によって内包が非関係属性と関係属性をともに有する概念束を構成する.

- ・関係概念解析は知識表現のオントロジー概念として表現できる関係構造を提供する.

1) 関係の単純な取込み

関係概念解析は対象と属性の関係を記述する形式文脈と対象間の関係を与える関係文脈の集合が必要であり, これを関係概念族という. 形式定義はスライド 29 のとおりである. 単純な関係概念族の例をスライド 30 に示す. この関係概念族は二つの形式文脈とひとつの関係文脈からなる.

形式概念と関係概念を統一的に扱うひとつの方法は形式文脈(ソース)に関係文脈を付加することである. これをスライド 30 の例で説明するなら, 形式文脈 $K_{動物}$ に関係文脈 $R_{棲む}$ をスライド 31 のように並置することである. この文脈を $K_{動物} + R_{棲む}$ で表わすことにしよう. $K_{動物} + R_{棲む}$ の Hasse 図を描くとスライド 31 の右図になる. しかし, この方法は陸生あるいは水生の動物は何かという問いに答えることができない. これを解決するには, 関係文脈から生成される関係を “関係

関係概念族(RCF)

- ・ RCF $F = (K, R)$
 K : 形式文脈 $K_i = (O_i, A_i, I_i)$ の集合
 R : 関係文脈 $R_j = (O_k, O_l, I_j)$ の集合
 (O_k, O_l) は文脈 $(K_k, K_l) \in K^2$ の対象集合 $I_j \subseteq O_k \times O_l$
 K_k : ソース, K_l : ターゲット
 $K_k = K_l$ でもよい

関係概念族の例:

- ・ ビザ・データ
 〈対象×属性〉
 $K_{人} = \text{人ID} \times \text{氏名}$, $K_{ビザ} = \text{ビザID} \times \text{ビザ名}$,
 $K_{食材} = \text{食材ID} \times \text{食材名}$, $K_{国} = \text{国ID} \times \text{国名}$
 〈対象×対象〉
 $R_{食べる} = \text{人ID} \times \text{ビザID}$, $R_{食む} = \text{ビザID} \times \text{食材ID}$
 $R_{生産国} = \text{食材ID} \times \text{国ID}$, $R_{国籍} = \text{国ID} \times \text{人ID}$

概念”として形式概念に取り込まなければならない。そのために Rlive で表現される“棲む”場所を形式文脈 Kplaces の概念の中で読み替える操作，たとえば“川”は概念 c7 (陸生，{川，海})に現れるので概念 c7 で置き換えを行う。同様に“サバンナ”，“森”は概念 c5 で置き換える (スライド 32)。この操作によって関係を取り込んだ概念束を構成するこ

とができる (スライド 33)。

2) 適用例—引用関係の逐次処理

主題別に分類されたクロス表 (スライド 33) があるとその概念図はスライド 34 になる。一方論文間の引用関係がスライド 35 で与えられているとき，主題別概念図に引用関係を取り込むことができる。引用関係は順次，

関係概念の処理: 単純な例

Kanimal	草食	肉食	雑食
熊			X
ライオン		X	
ワニ			X

$K_{動物} = 動物 \times 食性$
 $K_{棲息地} = 場所 \times 水/陸生$
 $R_{棲む} = 動物 \times 場所$

Kplaces	水生	陸生
川	X	
海	X	
サバンナ		X
森		X

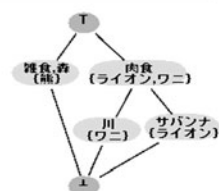
Rlive	川	海	サバンナ	森
熊				X
ライオン			X	
ワニ	X			

2012/8/28

知の創成シンポジウム

30

関係の単純処理



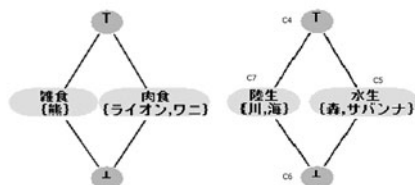
Kanmpic	草食	肉食	雑食	川	海	サバンナ	森
熊			X				X
ライオン		X				X	
ワニ		X		X			

2012/8/28

知の創成シンポジウム

31

関係属性としての取り込み



2012/8/28

知の創成シンポジウム

32

関係属性としての取り込み

Rlive1	草食	肉食	雑食	c4 (棲息地)	c5 (水生)	c6 (陸生)
熊			X	X		X
ライオン		X		X		X
ワニ		X		X	X	



2012/8/28

知の創成シンポジウム

33

論文の主題

ra: 要求分析

ad: アーキテクチャ設計

dd: 詳細設計

sm: ソフトウェア保守

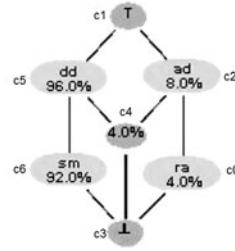
	ra	ad	dd	sm
maining				
fluc05			X	
fluc053			X	
fluc055			X	
fluc056			X	
fluc057			X	
fluc058			X	
fluc059			X	
fluc060			X	
fluc061			X	
fluc062			X	
fluc063			X	
fluc064			X	
fluc065			X	
fluc066			X	
fluc067			X	
fluc068			X	
fluc069			X	
fluc070			X	
fluc071			X	
fluc072			X	
fluc073			X	
fluc074			X	
fluc075			X	
fluc076			X	
fluc077			X	
fluc078			X	
fluc079			X	
fluc080			X	
fluc081			X	
fluc082			X	
fluc083			X	
fluc084			X	
fluc085			X	
fluc086			X	
fluc087			X	
fluc088			X	
fluc089			X	
fluc090			X	
fluc091			X	
fluc092			X	
fluc093			X	
fluc094			X	
fluc095			X	

2012/8/28

知の創成シンポジウム

33

主題別のHasse図



2012/8/28

知の創成シンポジウム

34

FCAの拡張

I

- 属性集合 (A) を属性集合のべき集合 (部分集合の全体, 2^A) で置き換える。
- 関係 I を $O \rightarrow 2^A$ への写像で置き換える。
- 以上の拡張によって, FCAの情報は失われず, 何も付け加えられず, FCAと同等である。

2012/8/28

知の創成シンポジウム

47

 2^A と論理式の関係

例1	二足	四足	肉食
ニワトリ	X		
ブタ		X	
ワシ	X		X
トラ		X	X

$M(\text{二足}) = \{\text{ニワトリ, ワシ}\}$
 $M(\text{四足}) = \{\text{ブタ, トラ}\}$
 $M(\text{肉食}) = \{\text{ワシ, トラ}\}$
 $M(\text{二足} \vee \text{四足})$
 $= M(\text{二足}) \cup M(\text{四足})$
 $= \{\text{ニワトリ, ブタ, ワシ, トラ}\}$
 $M(\text{肉食}) \wedge (\text{二足} \vee \text{四足})$
 $= M(\text{肉食}) \cap M(\text{二足} \vee \text{四足})$
 $= \{\text{ワシ, トラ}\}$

2012/8/28

知の創成シンポジウム

49

FCAの拡張

II

(O, A, I)	$(O, 2^A, i)$
	$2^A: A$ の部分集合
	$i: O \rightarrow 2^A$
	$i(o) := \{a \in A : \forall o \in O, (o, a) \in I\}$
$(o, a) \in I$	$i(o) \ni a$
$O' = \bigcap_{o \in O} i(o)$	
$A' = \{o \in O : i(o) \supseteq A\}$	

FCA $\equiv \langle 2^A, \supseteq \rangle$ とみなすことができるから, この 2^A を論理式の集合 L に置き換えてさらに一般化することができる。

2012/8/28

知の創成シンポジウム

48

 2^A の命題論理式への変換

命題 P に対応する集合を $M(P)$ と記し, 以下のよう
に再帰的に定義する:

$M(a) = \{A \in 2^A \mid a \in A\} \quad \forall a \in A$
 $M(0) = \emptyset \quad 0 \equiv \text{偽}$
 $M(1) = 2^A \quad 1 \equiv \text{真}$
 $M(P \wedge Q) = M(P) \cap M(Q)$
 $M(P \vee Q) = M(P) \cup M(Q)$
 $M(\neg P) = 2^A \setminus M(P)$

2012/8/28

知の創成シンポジウム

50

$$(o, a) \in I \Leftrightarrow i(o) = \{a\}$$

$$O' = \bigcap_{o \in O} i(o)$$

$$A' = \{o \in O \mid i(o) \supseteq A\}$$

である。

FCA は $\langle 2^A, \supseteq \rangle$ 上の \cap , \cup をそれぞれ最小上限, 最大下限操作とする束と同等である。このとき, 2^A は \supseteq を演繹関係, \cap を連言演算子, \cup を選言演算子とする論理とみなすことができる。命題論理の命題にスライド 50 に示すように 2^A 上の集合を対応させれば, 命題論理式に翻訳することができる。

2) 論理概念解析の定義

上で述べたように FCA の属性集合を命題論理の論理式で置き換えることができる。LCA ではこの論理をさらに一般化し, 任意の論理を選択することができる。論理の定義を正確にあたえつつぎのとおりである:
定義[論理] 論理とは 6 組 $(L, \sqsubseteq, \sqcap, \sqcup,$

$\top, \perp)$ のことである。ここで各記号は以下のとおり

- L : 論理式の言語
- \sqsubseteq : 従属関係
- \sqcap, \sqcup : それぞれ連言, 選言演算子
- \top, \perp : それぞれ恒真, 恒偽命題

この論理は束を構成する。

上の論理を用いて論理文脈の定義をスライド 52 に示した。外延・内包も同様に定義できるから LCA 上の概念とその順序も定義される。したがって, 概念集合は概念束になり, Hasse 図で表すことができる (スライド 55)。

3) 論理概念解析の適用

論理概念解析の適用は難しいことではない。簡単な例でいえば, 与えられたクロス表を簡単化するために複数の属性を連言で結合し, その組み合わせを属性とする操作も論理

命題論理の演繹関係

命題論理の演繹関係をつぎのように定義する:

$$P \vdash Q \Leftrightarrow M(P) \subseteq M(Q)$$

例: 肉食 \vdash 二足 \vee 四足

$\therefore M(\text{肉食}) = \{\text{ワシ, トラ}\}$

$M(\text{二足} \vee \text{四足}) = \{\text{ニワトリ, プタ, ワシ, トラ}\}$

$M(\text{肉食}) \subseteq M(\text{二足} \vee \text{四足})$

肉食 \vdash 二足 \vee 四足

2012/8/28

知の創成シンポジウム

51

論理概念解析: 文脈

I

- 文脈 (O, L, i)
- O : 対象の集合
- $\langle L; \vdash \rangle$ は論理式の束
- $i: O \rightarrow L$ への写像
- $X \subseteq O, f \in L$ に対し,
- $' : 2^O \rightarrow L, O' := \bigvee_{o \in O} i(o)$
- $' : L \rightarrow 2^O, f' := \{o \in O \mid i(o) \vdash f\}$

対象	記述
x	a
y	b
z	$c \wedge (a \vee b)$

2012/8/28

知の創成シンポジウム

52

論理概念解析: 概念と概念束 II

• 文脈 (O, L, i) において概念は対 $c = (X, f)$ である:

$X' = f, f' = X$ であるような $X \subseteq O, f \in L$ である。

対象の集合 X を外延, f をその内包という。

FCA との相違点: 内包が L の論理式である。

• $(X_1, f_1), (X_2, f_2) \in \mathcal{C}(O, L, i)$ の順序 \leq^c を

$(X_1, f_1) \leq^c (X_2, f_2) \iff \text{定義 } X_1 \subseteq X_2$

と定義する。

2012/8/28

知の創成シンポジウム

53

論理概念解析: 例

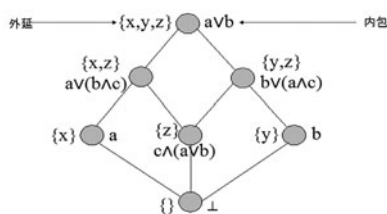
対象	記述	対象	記述
x	a	ニワトリ	二足
y	b	プタ	四足
z	$c \wedge (a \vee b)$	{ワシ, トラ}	肉食 \wedge (二足 \vee 四足)

2012/8/28

知の創成シンポジウム

54

論理概念解析: 概念束例



2012/8/28

知の創成シンポジウム

55

概念解析である。この種の操作をホーソン実験の分析に適用してみよう。表3のクロス表の2つの属性、学歴と勤続年数をひとつにまとめるため学歴 \wedge 勤続年数としよう。学歴 = 小卒 \vee 中卒 \vee 専卒, 勤続年数 = 短い \vee 中 \vee 長いであるから, 学歴 \wedge 勤続年数 = (小卒 \wedge 短い) \vee (小卒 \wedge 中) $\vee \dots \vee$ (専卒 \wedge 長い) である。いま対象を工員 W1 ~ W9 に限れば, 与えら

れたクロス表の学歴 \wedge 勤続年数の組み合わせは (小卒 \wedge 短い) \vee (中卒 \wedge 短い) \vee (中卒 \wedge 中) しかない。つまり二つの属性、学歴と勤続年数を連言すれば、このデータでは3個のエントリで済むことになる。このような操作をしたクロス表から描いた概念図を以下に示す。

図3に比べ概念図が簡略されていることがわかる。このような論理概念操作は、Web 検索時に無意識に行っていることに似ている。検索結果が多いとき、件数を少なくしなおかつ目的の情報に絞込む操作は論理概念解析そのもののなのである。

4. まとめと結論

社会情報解析への適用・応用を意識しながら FCA/RCA/LCA について述べてきた。FCA は対象に固有の属性から形式概念を見出し、概念間の関係を表現する数理的方法で

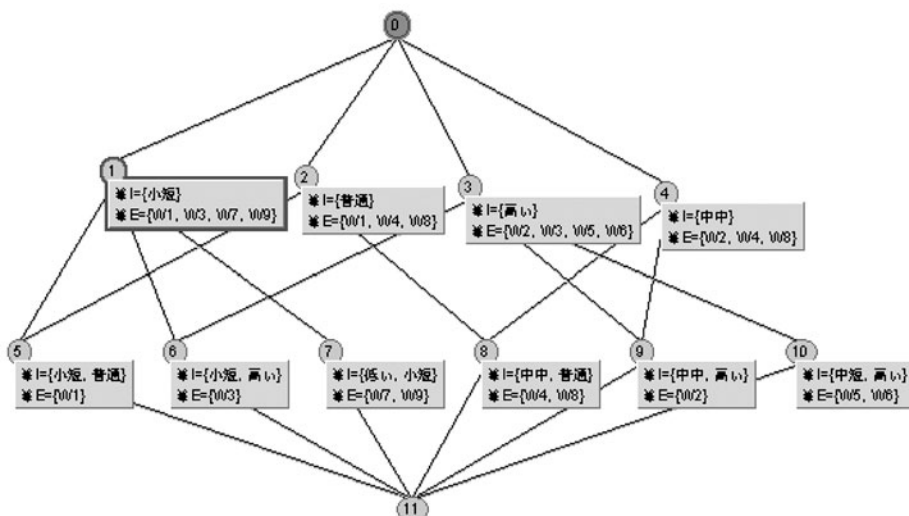


図5 学歴・勤続年数をひとまとめにした概念束

ある。RCAは対象間の関係に関係属性として捉えることにより形式概念へ取り込むことを可能にする。この結果、対象×属性から形成される概念構造をさらに多様に捉えることができる。LCAは、属性を論理式に拡張することによってより一般的な概念間の関係を表現することができる。しかも、これら三つは同一の考え方で統一されているのでその全体像が把握しやすいといえる。

FCA/RCA/LCAの研究はここで強調したような社会情報解析への応用に尽きるものではない。むしろ、知識発見、データマイニング、オブジェクト指向プログラミングに基づくソフトウェア工学、オントロジー工学、セマンティック工学など情報科学での研究が活発に行われている。

最後に、FCA/RCA/LCAによるアプローチは数理的基礎が明確であり、その結果は概念的思考を助け全体的見通しを与える方法であることを改めて強調するとともに、日本で

はFCA, RCA, LCAへの関心が極めて低いのでひとりでも多くの研究者が関心をもって下さることを願いながら筆を擱く。

謝辞

本稿の図1～図5は形式・関係概念解析のオープンプラットフォーム Galicia-v2.0を用いて作成した。記して開発者・関係者各位に謝意を表する。

参考文献

- Homans, G. C. (1950) The Human Group, New York: Harcourt. (=1959, 馬場明男・早川浩一訳, 『ヒューマン・グループ』誠信書房)
- 長田博泰(2004)「形式概念にもとづく質的分析」, 『社会情報』(札幌学院大学社会情報学部紀要), Vol.4 No.1, pp.19-37
- 進藤勝美(1978)『ホーソン・リサーチと人間関係論』産業能率短期大学出版部