

形式概念を用いたデータ解析: 幼児発達行動の分析

Data Analysis Using Formal Concept Analysis—an analysis of developmental behavior data of children

長田 博泰
NAGATA Hiroyasu

Formal Concept Analysis(FCA), introduced as a formalization of the concept of ‘concept’, has grown to a powerful theory for data analysis, knowledge discovery, and ontology engineering. This paper describes a method based on the use of FCA for the data analysis when dealing with real-world data sets. The method focuses on concepts that are definable by conjunctions of attributes, and make it possible not only to classify data based on formal concepts, but also to clarify dependency relations among attributes. The usefulness and clarity of the method are illustrated by data analysis of developmental behavior of children.

キーワード：形式概念解析，質的研究，データ解析，発達行動

Keywords : Forma Concept Analysis , Qualitative Research , Data Analysis, Developmental Behavior

1 はじめに

人文・社会科学の研究では、一つあるいは少数事例の徹底的な調査や綿密な分析に基づき立論する、いわゆる質的研究（質的分析）は有効な研究方法の一つである。質的研究はつぎの特徴を有する。すなわち、(1)事例一つ一つを全体としてとらえようとするので、各事例を変数に分解できないものとみる。(2)社会現象の多様性や因果関係の複雑さを重視するため、一般的傾向からずれた事例を例外や誤差として処理しない。(3)質的説明を多用して、各事例に見られる固有な現象の理解と説明をめざす（鹿又，2001：3）。

しかし、このアプローチは分析結果の一般性や客観性を保証しないという指摘をうけやすい。それは研究者の判断にゆだねられる部分が多く、また事例の選択が典型事例や極端な事例にかたよりがちであり、少数の事例、例証にとどまることも少なくないためである。

Ragin(1984)は、質的研究において現象の包括的説明力を高めるため、その要因の論理的関連を見出すブール代数的方法を提案した。

本稿はこの方向をさらに拡張し、より多様な分析を可能にする方法として、1980年代前半に提案された形式概念解析を採用する(Wille, 1982)。この方法はつぎの点でブール代数的方法の拡張とみなすことができる。(1)変数値は二値に限らず、任意の名義尺度、順序尺度を扱うことが可能である。(2)現象の多様性をそのまま表現することができる。(3)分析手続きが客観的である。(4)より節約的なモデルを想定しない(結果としてより節約的なモデルが導出される可能性はある)。

長田(2004)は、この方向に沿って多様な価値判断・評価を記述するために形式概念束のHasse図を援用し、その比較分析を行った。さらに長田(2006)では、この方法が社会調査データにも適用可能であることを示すために「敬語意識に関する調査データ」を主として属性間に成立する関係に注目し分析したが、大量のデータの全体的構造を捉える形式概念解析の手法に欠けていた。

本稿はこの欠を補うためにつぎの2つを目的とする。データ量が多い場合にもより一層適

用しやすいよう形式概念解析の方法を拡張すること、および拡張した方法を用い幼児の発達行動記録を分析し、発達行動パターンの全体的構造および行動間の関係、すなわち行動パターンを見出し、社会調査など様々なデータ解析方法としての有効性と簡明性を実証することである。

以下、2節で形式概念を用いたデータ解析の展開に必要な定義、分析方法およびその拡張について説明し、3節では具体的なデーター幼児の発達行動記録ーに適用し、全体的構造とその特徴および年(月)齢と性別による発達行動との関連を明らかにする。4節ではまとめと今後の課題を述べる。

2. 形式概念解析

形式概念解析 (Formal Concept Analysis, 以下 FCA と略す) は、1980 年代半ばに Wille らが開拓した、記号/カテゴリカルデータを順序関係^{註1)}に基づいて数学的に解析する方法である。FCA の基本的考え方はガロア結合とよばれる代数的関係を基に属性あるいは対象をグルーピングし、概念束とよばれる半順序集合を構成することである。このグルーピングは属性あるいは対象のどの集合が概念と呼ばれる一様なまとまりであるかを決定する。概念とは外延と内包を統一する哲学的基準である。概念の外延は概念に属する対象からなる集合である。他方、概念の内包は概念に属するすべての対象が共有する属性の集合である。本節では FCA の必要最小限の事項を説明する。詳細は Ganter&Wille(1999) , Davey & Priestley(2002)を参照されたい。

2.1 形式概念と概念束

動物	小	中	大	二足	四足	羽	毛	飛	狩	走	鬣	蹄
ハト	X			X		X		X				
ニワトリ	X			X		X						
フクロウ	X			X		X		X	X			
ワシ		X		X		X		X	X			
イヌ		X			X		X			X		
トラ			X		X		X		X	X		
ライオン			X		X		X		X	X	X	
ウマ			X		X		X			X	X	X
ウシ			X		X		X					X

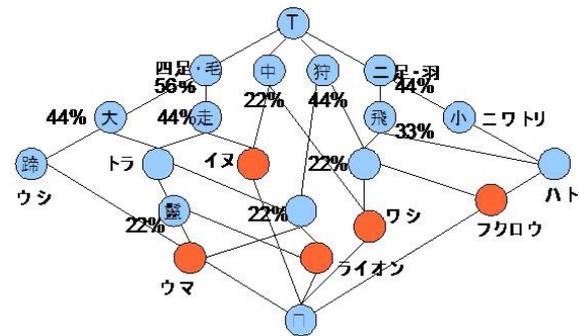


図1 概念束図

表1の簡単な例を取り上げる。表1の第1行を属性、左端の列を対象という。表中の“x”印は行の先頭の対象と列の先頭の属性の間に関係が成立していることを表す。例えば“ハト”は“小”さく、“二足”であり、“羽”があって“飛”ぶこと示す。

束(そく)図のノードは概念であり、属性あるいは対象の特別な種類の集合である。例えば、属性集合{二足, 羽}は概念である。なぜなら、この表から“二足, 羽”である対象は、集合{ハト, ニワトリ, フクロウ, ワシ}であり、逆に{ハト, ニワトリ, フクロウ, ワシ}に共通する属性集合は{二足, 羽}であるからである。一般に、概念XはXのすべての性質をもつすべての対象が共有する属性と一致するような属性集合によって与えられる、すなわちXはその外延の内包と一致する。しかし、{二足}は概念ではない、なぜならその外延{ハト, ニワトリ, フクロウ, ワシ}から内包を求めると、異なる集合{二足, 羽}になるからである。概念の集合が決まると、集合の包含関係によって概念を順序づけることができ、その関係を図1(これをHasse図という)のよう

に描くことができる。

以上をやや形式的に説明しよう。FCAでは形式文脈 (O, A, R) を考える、ここで O は対象の集合である、 A は属性の集合、そして R は O と A の間の二項関係であり、形式文脈はしばしば表（クロス表という）で表わされる。FCAはこの低レベルのクラス分けを高レベルの束構造に体系的に変換し、概念階層を図示化する。

表1の例では

$O = \{\text{ハト, ニワトリ, フクロウ, ワシ, イヌ, トラ, ライオン, ウマ, ウシ}\}$

$A = \{\text{小, 中, 大, 二足, 四足, 羽, 毛, 飛狩, 走, 鬣, 蹄}\}$

である。表中の”x”印は、行の先頭の対象と列の先頭の属性間に関係が成立していることを示す。例えば、表は“ハト”が“小, 二足, 羽, 飛”であることを示す。表を読み取って、二項関係をつぎのように表現することができる。

$R = \{(\text{ハト, 小}), (\text{ハト, 二足}), (\text{ハト, 羽}), (\text{ハト, 飛}), (\text{ニワトリ, 小}), (\text{ニワトリ, 二足}), (\text{ニワトリ, 羽}), \dots, (\text{ウシ, 大}), (\text{ウシ, 四足}), (\text{ウシ, 毛}), (\text{ウシ, 蹄})\}$

文脈 (O, A, R) に関して任意の $X \subseteq O, Y \subseteq A$ に対し内包, 外延をそれぞれつぎのように定義する。

$(X)' = \{a \mid \forall x \in X, (x, a) \in R\}$ (内包)

$(Y)' = \{o \mid \forall y \in Y, (o, y) \in R\}$ (外延)

概念とは、 X の内包が Y であり Y の外延が X であるような対 (X, Y) である、すなわち (X, Y) が概念である $\Leftrightarrow (X)' = Y$ かつ $(Y)' = X$

文脈 (O, A, R) に関し、対 $((), ())$ はガロア結合である、すなわち任意の $X \subseteq O, Y \subseteq A$ に対し

$(X)' \supseteq Y \Leftrightarrow X \subseteq (Y)'$

対 $(U, V), (X, Y)$ の順序関係を以下のように定義する。

$(U, V) \leq (X, Y) \Leftrightarrow U \subseteq X$ かつ $V \subseteq Y$

(O, A, R) の概念集合はこの順序関係に関して完備束^{注2)}になる。

2.2 含意規則

属性表とその束は同一情報を異なる視点から捉えたものであり、目的に応じて使い分けることができる。さらに含意規則と呼ばれる別な視点からも捉えることができる。 $A, B \subseteq A$ を2つの属性集合とする。 A の属性をもつ対象がまた B の属性を持つときに限って A は B を含意するという、すなわち

$A \rightarrow B \Leftrightarrow$ すべての $o \in O$: (すべての $a \in A$: $(o, a) \in R$) \Rightarrow (すべての $a \in A$: $(o, a) \in R$)

$B = \{b_1, \dots, b_k\}$ に対し、すべての b_i について $A \rightarrow b_i$

のときに限って $A \rightarrow B$ が成り立つ。属性間の含意は、直感的には、束を上の方にたどって見出すことができる。表1で成立する含意規則を表2に掲げる。四足 \rightarrow 毛；毛 \rightarrow 四足；二足 \rightarrow 羽；羽 \rightarrow 二足；走 \rightarrow 四足, 毛；大 \rightarrow 四足, 毛などが成立している。なお、改めて指摘することでもないが、含意関係は因果関係を表現するものではないことに留意されたい。含意関係が因果的關係であることを示すにはさらに新たな考察が必要である。

2.3 粗い概念束と連関規則

形式概念解析を現実の問題に適用する場合、つぎの2つの問題が発生する。一つは対象数、

件数	含意規則	
5	{毛}	\rightarrow {四足}
5	{四足}	\rightarrow {毛}
4	{走}	\rightarrow {四足, 毛}
4	{羽}	\rightarrow {二足}
4	{二足}	\rightarrow {羽}
4	{大}	\rightarrow {四足, 毛}
3	{飛}	\rightarrow {二足, 羽}
3	{小}	\rightarrow {二足, 羽}
2	{蹄}	\rightarrow {大, 四足, 毛}
2	{鬣}	\rightarrow {大, 四足, 毛, 走}
2	{四足, 毛, 狩}	\rightarrow {大, 走}
2	{二足, 羽, 狩}	\rightarrow {飛}

属性数が多いと膨大なコンピュータパワー

を必要とすることである。いまひとつは、概念数が多くなると概念束を事実上、描画することができなくなり、たとえ描けたとしても全体像を把握することが著しく困難になることである。

そこで実際のデータに適用するに当たってつぎの方法を導入する。概念束の全体構造の概略と特徴を把握するために”粗い概念束“を描くことである(Stumme, 2001, 2002)。また、対象が複雑、多様な場合、含意規則の個数も膨大になり、少数の規則によって全体を把握することはできない場合が少なくない。この点を考慮し、粗い概念束の導入に合せ、含意規則を弱めた連関規則を用いることにする。

概念 (X, Y) の支持度を外延 X の構成比として定義する。すなわち

$|X|/N$ ここで $|X|$ は X の要素数、N は対象数 = $|O|$ である。

与えられた支持度 ($\in [0, 1]$) 以上の概念だけからなる集合は、最小要素を付け加えると概念束になる(Stumme, 2002:194)。この支持度を最小支持度、束を”粗い概念束”^{注3)}という。なお、しばしば最小要素を省略するが、その場合も“粗い概念図”と呼ぶことにする。

概念束の一部の概念しか含んでいない“粗い概念束”を記述するためにデータマイニング分野における有力な方法として用いられている連関規則を採用する。上述の含意規則が確実な規則(信頼度 100%あるいは 1)なのに対し、連関規則その信頼度を弱めて対象を記述しようとするものである。連関規則を厳密に定義しよう(Pang-Ning Tan, 2006:330)。

連関規則とは、 $X \rightarrow Y$ である、ただし、 $X, Y \subseteq A$ で、 $X \cap Y = \emptyset$ 。連関規則の支持度をつぎのように定義する。

支持度 $s(X \rightarrow Y) = |(X \cup Y)'| / N$

また、信頼度を次式で与える。

信頼度 $c(X \rightarrow Y) = |(X \cup Y)'| / |X'|$

図1の概念束を最小支持度 0.4, 0.3, 信頼度 0.6 で描いた粗い概念図を図2に示す。概念

束の特徴が捉えられ、支持度を小さくすると次第に概念束の仔細が見えてくる。すなわち二足・羽と四足・毛に大きく分かれ、各々に飛走、および大きさ属性、小、大が加わっている。狩はまだどこにも結合していないが、さらに詳しく描くと共通な属性であることがわかる。この方法によって概念的クラスタリングを行うことができる。

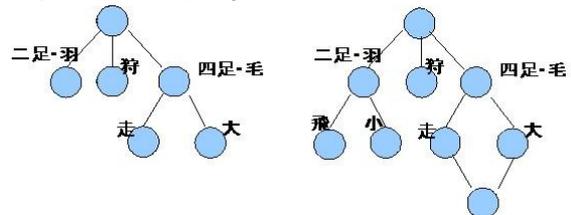


図2 図1の粗い概念束

3. 適用—幼児発達行動の分析

以上の定義および方法を準備として具体的データ、ここでは幼児の発達行動記録に形式概念を適用し、乳幼児の行動の特徴等を分析しよう。データは、京都市法光院保育所の乳幼児の9つの行動に注目し、ひと月ごとに一年間にわたって観察記録したものである。データの一部を表3に示す。そのようなデータが7年分ある。各年(平成11~平成18年)の4月における延べ96人(男44人、女52人)に観察された行動を分析対象とする。

表3 発達行動記録(一部)

対象	性別		月齢	観察行動									
	男	女		床に伏して泣く	食事の時泣く	噛みつく	ひひかく	たたたく	突き飛ばす	髪を引っぱる	物を奪う	物を投げる	
c1		x	19	x				x				x	x
c2	x		19					x				x	x
c3		x	18			x		x	x			x	x
c4		x	16					x	x			x	x
c5		x	16					x	x			x	x
c6		x	15		x		x	x				x	x
c7		x	14		x			x			x	x	x
c8	x		13		x							x	

このデータをつぎの2つの観点から分析する。ひとつは、観察対象とした9つの行動の全

体的構造とその特徴および行動間の関係、すなわち行動間に成立する含意関係を闡明することである。いまひとつは、乳幼児の月齢と性別による発達行動の全体的特徴および月齢と行動の関係を明らかにすることである。

3.1 発達行動の全体的構造とその特徴

月齢、性別を考慮せず、96人の乳幼児を9行動だけをそのまま形式概念解析すると、ノード数（概念数）122個の概念束が得られる。図は煩雑なので省略するが、この分析からつぎのことがわかる。観察した9つの行動のうち「髪を引っ張る」を除く8つの行動が最上位ノード（“T”）の直下の概念—以下、基本行動とよぶ—として出現する。これはこれら8つの

基本行動のそれぞれの外延が互いに他に含まれることはないことを意味する。では、「髪を引っ張る」はどうなっているのだろうか。表4に掲げる含意関係がこれを示す。すなわち、「髪を引っ張る」は「物を奪う」に吸収されるのである。これらの含意関係は行動間のパターンを簡潔に捉えているが、支持度が示すように成立する範囲は比較的小さい。最も高い含意規則 {噛む, 物投げる} → {たたく, 物奪う} でも、96人中14人、支持度0.15程度であり、全体の一部の特徴しか表わしていない。

このままでは全体的構造が把握できないので、粗い概念束を描く。最小支持度0.5, 0.4, 0.3, 0.2で描いた粗い概念束の結果がそれぞれ図3 a, b, c, dである。

表4 全体の含意規則

人数	支持度	含意規則	
14	0.15	{噛む, 物投}	→ {たたく, 物奪}
14	0.15	{ひっかく, 突飛}	→ {たたく}
13	0.14	{髪引}	→ {物奪}
10	0.10	{噛む, 突飛}	→ {たたく, 物奪}
9	0.09	{突飛, 髪引, 物奪, 物投}	→ {たたく}
9	0.09	{食泣, 突飛}	→ {物奪}
9	0.09	{床泣, 突飛}	→ {物奪, 物投}
9	0.09	{噛む, ひっかく}	→ {たたく, 物奪}
7	0.07	{床泣, ひっかく}	→ {物奪, 物投}
5	0.05	{食泣, ひっかく}	→ {物奪, 物投}
5	0.05	{床泣, 食泣, たたく, 物奪}	→ {物投}
5	0.05	{噛む, ひっかく, たたく, 髪引, 物奪, 物投}	→ {突飛}
5	0.05	{噛む, たたく, 突飛, 髪引, 物奪, 物投}	→ {ひっかく}
4	0.04	{床泣, 噛む}	→ {たたく, 物奪, 物投}
3	0.03	{食泣, 噛む, たたく}	→ {物奪, 物投}
2	0.02	{食泣, たたく, 突飛, 髪引, 物奪, 物投}	→ {ひっかく}

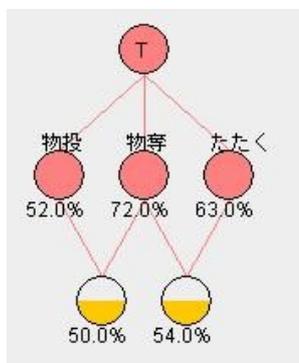


図3 a 粗い概念束(最小支持度 0.5)

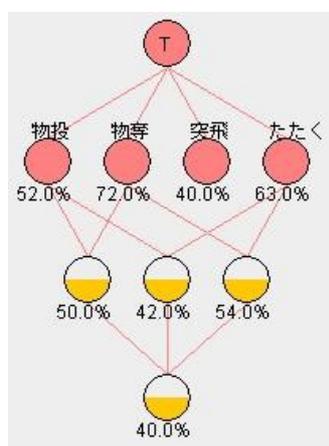


図 3 b 粗い概念束(0.4)

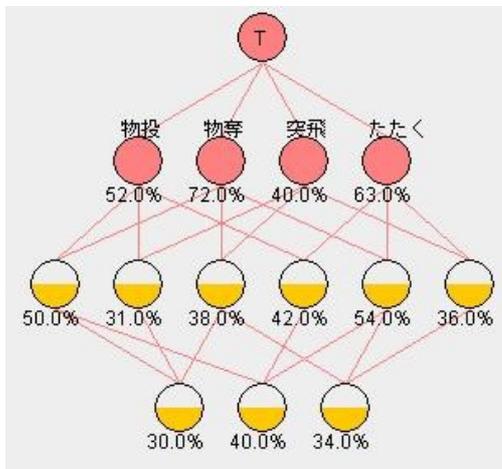


図 3 c 粗い概念束(0.3)

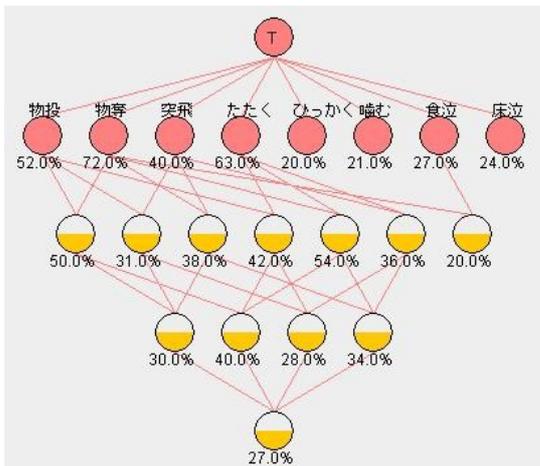


図 3 d 粗い概念図(0.2)

図 3 a-d には行動の全体的構造が表現され、つぎの事実を読み取ることができる。

- 1) 50%以上観察されたのは3つの基本行動は「物を奪う」、「たたく」、「物を投げる」とその組合せパターン「物を奪う・物を投げる」、「物を奪う・たたく」である。
- 2) 最小支持度を0.4にすると「物を奪う」、「たたく」、「物を投げる」の3基本行動に新たに「突き飛ばす」が加わり、これらの組合せが増える。
- 3) 最小支持度を0.3では新たに加わる基本行動はなく、2)の4つの行動の組合せパターンが増すだけである。
- 4) 最小支持度0.2以上0.3未満の基本行動は「ひっかく」、「床に伏してなく」、「食事の時泣く」、「噛む」の4つである。
- 5) 最小支持度0.2以上の概念数は概念全体(122ノード)の5分の1程度であり、多様な行動パターンは支持度0.2未満の概念に現れる。
- 6) 「髪を引っ張る」は3-dに現れていないので、その支持度は0.2未満である。

表 5 月齢3分類の人数(構成比)

性別	月齢13～21ヶ月	月齢22～26ヶ月	月齢27～33ヶ月	合計
男	15 (48)	17 (52)	12 (37.5)	44 (46)
女	16 (52)	16 (48)	32 (62.5)	52 (54)
合計	31 (32)	33 (34)	32 (33)	96 (100)

以上の結果は9つの行動間の関連を表面的に捉えているにしても、乳幼児の発達と行動の関係が見えてこないため、つぎにデータとして与えられている月齢と性別を考慮し、形式概念解析を適用する。

3.2 月齢、性別と行動の関係

月齢、性別を属性として扱い、概念束の上にこれらが属性として表れるような分析

も可能であるが、少しでも見通しがよい概念束図を描くため、与えられたデータを表5に示すように予め月齢別、性別に分類し、行動だけを属性として扱うことにする。月齢の区切りとして対象データをほぼ三分する13ヶ月～22ヶ月、23ヶ月～26ヶ月、27ヶ月～33ヶ月を選択した。その構成を表5に示す。

月齢3区分×男女別の6ケースを形式概

念解析した結果の概念束（図4 a-c, 5 a-c）と含意規則（表6, 7）を示す。これらの6つの図から月齢および性別による9つ

の行動に関し以下の顕著な特徴を読み取ることができる。

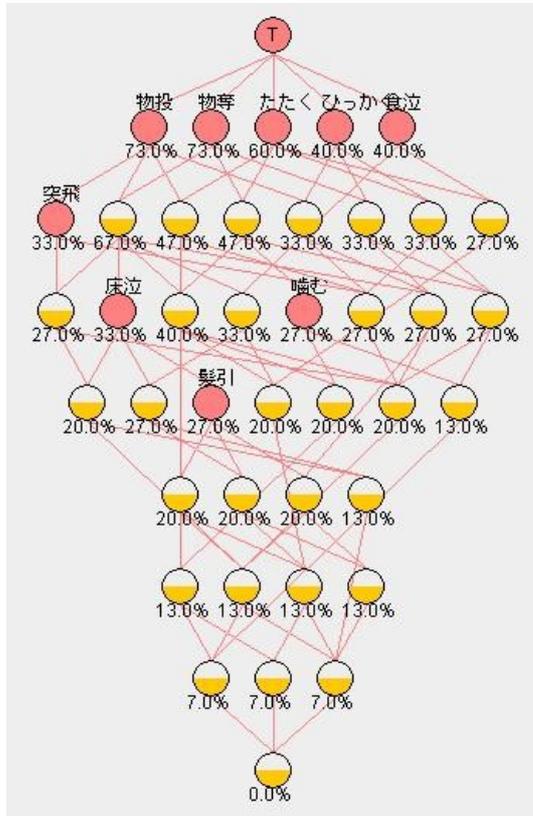


図4-a 月齢13～21ヶ月，男

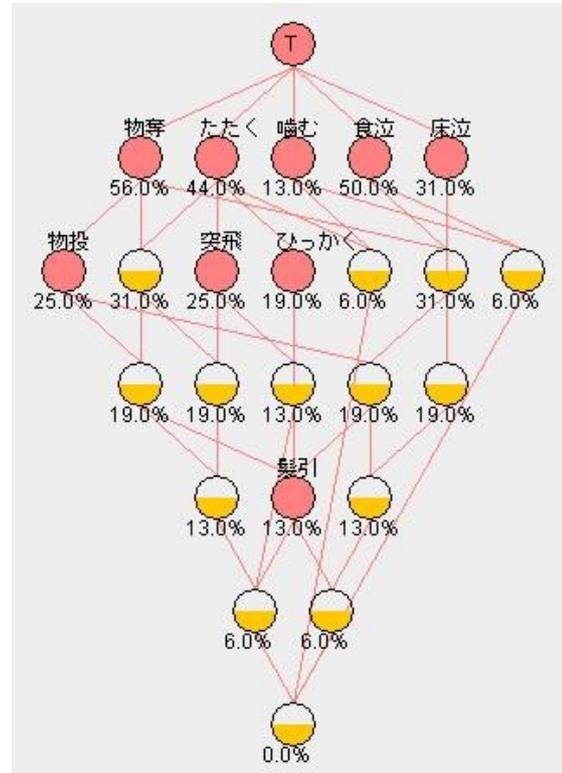


図5-a 月齢13～21ヶ月，女

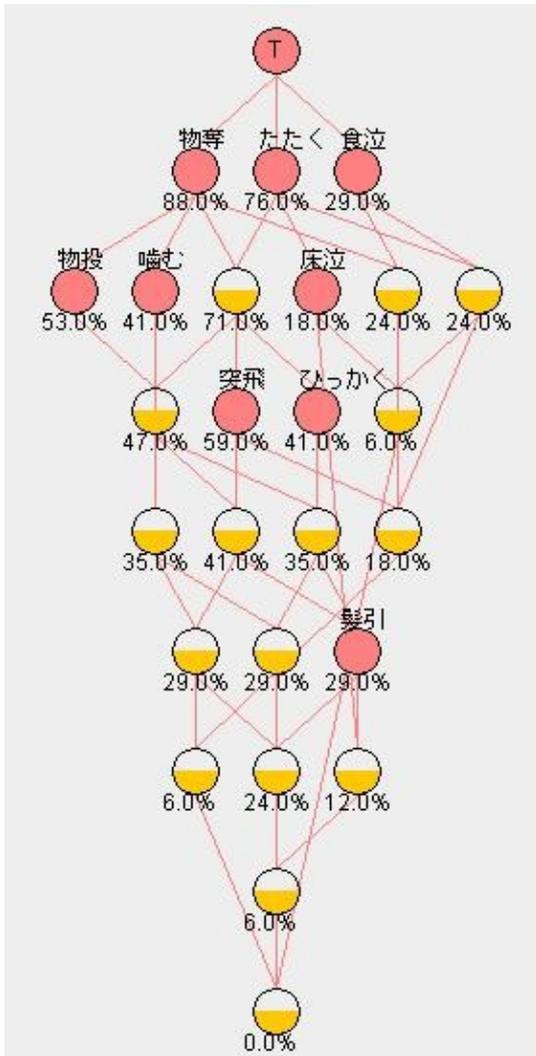


図4-b 月齢22ヶ月～26ヶ月，男

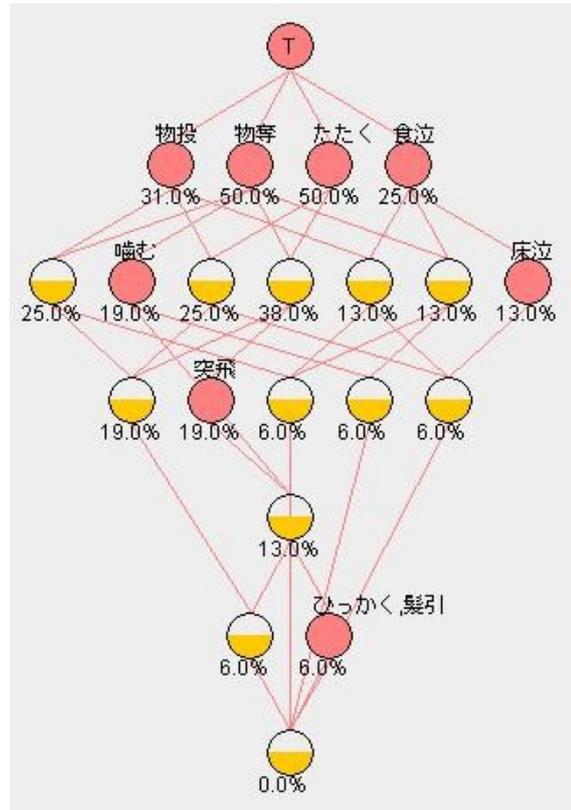


図5-b 月齢22ヶ月～26ヶ月，女

表6 月齢27～33ヶ月，男			
人数	含意規則		
9	{物投}	→	{物奪}
8	{たたく}	→	{物奪}
6	{突飛}	→	{物奪}
3	{噛む}	→	{たたく，物奪，物投}
2	{床泣}	→	{物奪}
1	{床泣，物奪，物投}	→	{突飛}
表7 月齢27～33ヶ月，女			
人数	含意規則		
12	{物投}	→	{物奪}
10	{突飛}	→	{物奪，物投}
6	{床泣}	→	{たたく}
3	{食泣}	→	{たたく，突飛，物奪，物投}
2	{ひっかく}	→	{たたく，突飛，物奪，物投}
1	{髪引}	→	{たたく，突飛，物奪，物投}
1	{噛む}	→	{床泣，食泣，ひっかく，たたく，突飛，物奪，物投}

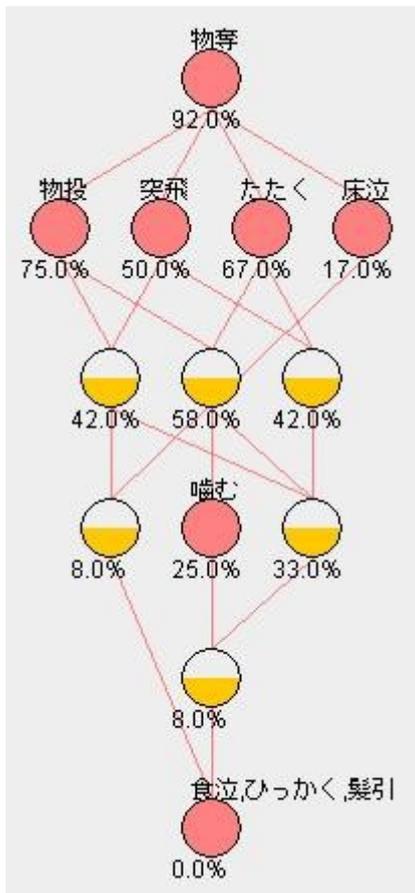


図4-c 月齢27～33ヶ月，男

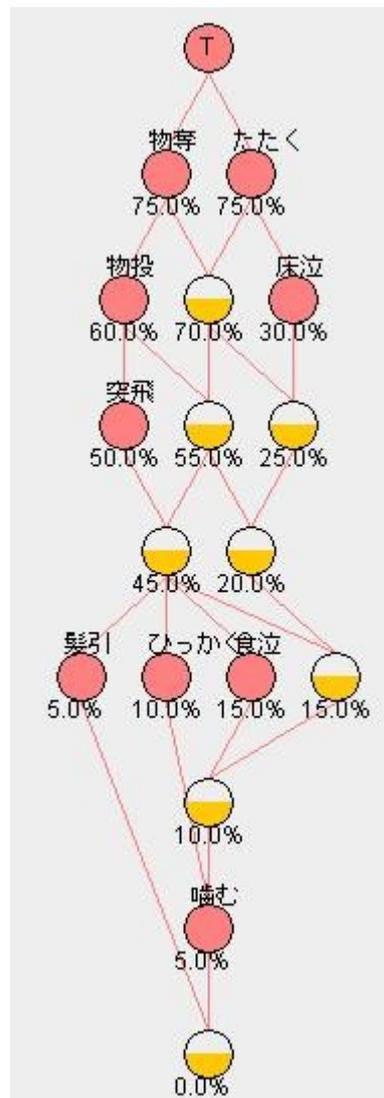


図5-c 月齢27～33ヶ月，女

- 1) 月齢の増加とともに、男女とも基本行動のタイプが少なくなり、男児では月齢とともに5, 3, 1種に減少してゆくのに対し、女児は5, 4, 2種であり、やや緩やかに減ってゆく。
- 2) 最終的（月齢27～33ヶ月）には、男女の基本行動はそれぞれ「物奪う」、「物奪う、たたく」に収束してゆく。しかし、図4-c, 5-cおよび表6, 7の含意規則からわかるようにそこに至る過程でやや複雑な行動パターンを示すものが少数認められる。
- 3) 月齢27～33ヶ月の概念束から9つの行動を「物奪う、物投げる、突き飛ばす、たたく、床に伏して泣く」と「髪を引っ張る、ひっかく、嘔む、食事時泣く」に二分できる。しかし、「床に伏して泣く」は他の4つの基本行動に比べ、比率が低いので、やや特異な傾向があると思われる。
- 4) この月齢分類からみると、月齢26ヶ月以下では「食事時泣く」が基本行動になっているが、月齢27～33ヶ月で急速に減ずるのに対し、「床に伏して泣く」の比率は高くないが、最後まである程度の比率で残る。

3.3 発達行動結果の考察

以上で得られた分析結果が幼児発達行動において含意するところを考察する。顕著に観察された「物奪う」、「たたく」、「突き飛ばす」などは比較行動学、人間行動学（ヒューマンエソロジー）でいうところの「攻撃」的行動とみなすことができるように思われる。そこで幼児の人間行動学的観点から発育行動の分類を試みる（Jones, 1972=1987）。

まず、9つの行動のうち「物奪う」、「たたく」を月齢27～33ヶ月の結果から「攻撃」的行動とみなすことに問題はないであろう。

「物投げる」、「髪を引っ張る」、「突き飛ばす」は「物を奪う」、「たたく」に含まれるか、もしくは関連性が強いので、「前攻撃」タイプとする。一方、残りの行動のうち「食事時泣く」は成長過程とともに減少する。しかし図4-c, 5-cにみられるように「床に伏して泣く」は少数であるが、月齢27～33ヶ月にも示す幼児がいる。両者は質を異にするとと思われるが、取りあえず両者を「不安」を表す行動と考える。残った「嘔む」、「ひっかく」は図4-c, 5-cが示すように4～5個の行動の組合せとして出現し、最終的に「攻撃1」基本行動と関連していることに注目すれば、タイプ1とみなすことも可能であるが、「不安」との関連も認められるので、別タイプ「攻撃2」としておく。以上の考察から9つの行動を以下のように分類することができる。

タイプ1（攻撃1）：「物奪う」、「たたく」

タイプ2（前攻撃）：「突き飛ばす」、
「物投げる」、「髪を引っ張る」

タイプ3（攻撃2）：「ひっかく」、「嘔む」

タイプ4（不安）：「食事時泣く」、
「床に伏して泣く」

9つの行動を以上のように分類すると、観察した乳幼児の発達行動をつぎのように概括することができる。「不安」の少ない幼児は「攻撃2」を経ずにタイプ1または2に至ると考えられる。他方、やや「不安」を感じる幼児は「攻撃2」行動を示しながら、タイプ1または2の行動パターンに落ち着くが、「不安」の強い子はこの「攻撃2」やとりわけ「床に伏して泣く」行動から抜けられないと考えられる。

5. おわりに

形式概念解析は、第一に概念の数学的表

現である、第二に数量的方法に拠らず、データを構造化・分析する数理的方法である、第三にデータとその属性の構造および含意関係を可視化する方法である。つまり、形式概念解析はデータを加工することなく、データの全体的構造を視覚化し、属性間の含意関係を明示的に捉えることを可能にする。さらに含意関係を利用し、属性間の論理的関係を見出すことができる。しかし、一般に形式概念解析の概念数が多くなると、たとえ概念束を描いたとしても全体像を把握することが著しく困難になる。このような場合、本稿で指摘したように”粗い概念束“を描くことが有効である。

形式概念解析の有効性と簡明性を示すため、乳幼児の発達と観察記録された「物を奪う、物を投げる、・・・」の9つの行動の関係を分析し、つぎの結果を見出した。月齢13～26ヶ月までは9つの行動が複雑に絡んでいるが、月齢27～33ヶ月では、多くの幼児が攻撃的行動「物を奪う、たたく」に収束するが、やや「不安」を示す幼児は「ひっかく、嘔む、食事時なく、床に伏して泣く」などの行動を示す。この経過を概念束として示すことができた。これは形式概念解析が質的分析にとって有効な方法であることを充分示すものである。しかし、ここで得られた分析結果を生かすには、保育現場でもう少し乳幼児の個別的観察を重ねる必要があるが、それは今後の課題である。

形式概念の方法が実際のデータに広く利用されるには、高速に概念束を計算する方法の研究と概念の関係が把握しやすく見通しのよい概念図を描くツールの開発が必須である。ここでは、概念束の計算および概念図の描画(図3-5)に試作中のツールを用いた。

今後、さらに見通しのよい概念図を描くツールを開発するとともに、人文・社会科

学におけるデータに形式概念解析を適用しその有効性を示してゆきたい。

注

1) O を集合 M における関係とする。 O が つぎの性質をもつとき、 O を順序関係という。

$$(1) \text{ 反射律 } xOx$$

$$(2) \text{ 推移律 } xOy, yOz \Rightarrow xOz$$

$$(3) \text{ 反対称律 } xOy, yOx \Rightarrow x=y$$

ある集合 M において一つの順序関係 O が与えられたとき、 (M, O) を順序集合という。たとえば、自然数の間の通常的大小関係 \leq は、自然数の集合 N における一つの順序関係であり、 (N, \leq) と記す。

2) 順序集合 L において、任意の二元 a, b に対して $\{a, b\}$ の上限および下限がいつも L の中に存在するとき、 L を束 (Lattice) という。順序集合 L の空でない任意の部分集合 A に対してその上限 $\sup A$ および下限 $\inf A$ が存在するとき、 L を完備束という。ここで、 $A (\subseteq L)$ の上限 (最小上界) $\sup A$ とはつぎの条件をみたす要素である：

$$(1) a \in A \Rightarrow a \leq \sup A$$

$$(2) (a \leq A \Rightarrow a \leq x) \Rightarrow \sup A \leq x$$

$\sup A$ と双対的に、すなわち上の二つの条件の \leq を \geq に置き換えて、 A の下限 (最大下界) $\inf A$ が定義される。

3) Stumme (2001, 2002) は“粗い概念束”を “Iceberg Concept Lattice” (直訳すれば、冰山(のような)概念束)と呼んでいるが、粗い概念束のほうが語感から想像しやすいと思われる。また、Stumme は最小支持度以上の概念だけを求め、描く方法を採用しているのに対し、ここではすべての概念を決定した後、概念束を描く際に最小支持度以上のノードのみを描く方法をとっている。このような違い

もあるので“粗い概念束”とよぶことにする。

謝辞

本稿で用いた幼児の「発達行動記録」データの利用を快諾してくださった京都法光院保育所の三好園長および保育士の方々、日頃議論し有益なコメントを頂いた上、データ提供仲介の労をとって下さいました田中一先生、および千葉正喜札幌学院大学教授にお礼申し上げます。匿名の査読者から有益な指摘とコメントを頂きました。記して謝意を表します。

引用文献

Davey & Priestley(2002): *Introduction to Lattices and Order*(2ndED.), Cambridge University Press

Ganter, B., Stumme, Wille(eds)(2005): *Formal Concept Analysis*, Lecture Notes in Computer Science 3626, Springer

Ganter, B and Wille, R. (1999): *Formal Concept Analysis*, Springer

Jones, N. B. (ed) (1972) : Ethological studies of child behaviour, Cambridge University Press=岡野恒也監訳 (1987) 『乳幼児のヒューマンエソロジー—発達心理学の新しいアプローチ』, ブレーン出版

鹿又伸夫・野宮大志郎・長谷川計二編著 (2001) : 『質的比較分析』ミネルヴァ書房

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. 14, No. 1, pp. 19-

37

(2005) : 形式概念にもとづく質的分析, 日本社会情報学会第10回大会報告要旨集, pp. 4-5

(2006) : 「社会情報解析への一寄与: 形式概念によるデータ解析」, 『社会情報』, Vol. 15, No. 2(103-118), 札幌学院大学社会情報学部

(2006) : 形式概念を用いたデータ解析, 日本社会情報学会第11回大会報告要旨集, pp. 105-106

Pang-Ning Tan, M. Steinbach and V. Kumar(2006) : *Introduction to Data Mining*, Pearson Education

Ragin, C. C., S. E. Mayer, and K. A. Drass (1984) : Assessing Discrimination : A

Boolean Approach, *American Sociological Review* Vol. 49 (April:221-234)

(1987=1993) : The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies. University of California Press.(鹿又伸夫監訳 『社会科学における比較研究—質的分析と計量的分析の統合に向けて』 ミネルヴァ書房)

Stumme, G. et al. (2001) : Conceptual Clustering with Iceberg Concept Lattices, *Proc. GI-Fachgruppentreffen Maschinelles Lernen '01*. October 2001

(2002) : Computing Iceberg Concept Lattices with TITANIC, *Journal on Data and Knowledge Engineering*, Vol. 42, issue 2, pp. 189-222

(2003) : Off to new shores: conceptual knowledge discovery and processing, *International Journal of Human-Computer Studies*, 59, pp. 287-325

- Wille, R. (1982): Restructuring lattice theory: an approach based on hierarchies of concepts, *Ordered sets* (ed. I. Rival) pp. 445-470, Reidel, Dordrecht-Boston
- Wolff, K. E. (1996): Comparison of graphical data analysis methods, *SoftStat '95 advances in statistical software* 5 (ed. Faulbaum, F. & Bandilla, W.) pp. 139 - 151, Lucius&Licius, Stuttgart