

# 因果帰納における有向・無向推論の切り替え：対称的否定と刺激の提示形式の影響

## Switching Between Directed and Undirected Reasoning in Causal Induction: Effects of Symmetrical Negation and Stimulus Presentation Format

富田 貴央<sup>†</sup>, 市野 弘人<sup>†</sup>, 青木 颯大<sup>†</sup>, 高橋 達二<sup>†</sup>, 樋口 滉規<sup>‡</sup>

Takao Tomita, Hiroto Ichino, Sota Aoki, Tatsuji Takahashi, Kohki Higuchi

<sup>†</sup> 東京電機大学, <sup>‡</sup> 中部大学

Tokyo Denki University, Chubu University

22rd133@ms.dendai.ac.jp

### 概要

本研究では、観察から因果関係の強さを推定する認知課題において、情報の提示形式（オンライン／サマリー）と否定の対称性（非対称／対称）が推論の形式に与える影響を検証した。実験参加者の平均回答値をクラスタリングし、それぞれ因果の方向性を考慮するモデル CS と考慮しないモデル UCS との適合を分析した結果、情報の提示形式や否定の対称性を操作することで因果の方向性を考慮した推論が促進されることが示唆された。

**キーワード：因果推論, 構造学習, 二重フレーム理論, Causal Support, Undirected Causal Support**

### 1. はじめに

因果帰納の研究において、人間が観察データに基づいて因果関係を帰納的に推論する際、課題特性や状況に応じて複数の推論形式を動的に使い分けられていることが指摘されている (Hattori et al., 2017)。また、これまでに人間の因果帰納を説明するため様々な数理モデルが提案されてきた。主な枠組みとしては、出来事の共起情報から因果強度を連続量で推定する「因果強度モデル」( $\Delta P$ , Causal Power など)と、因果関係が存在するという仮説の尤もらしさを推定する「因果構造モデル」(Causal Support (CS), Undirected Causal Support (UCS) など)が挙げられる (Griffiths & Tenenbaum, 2005)。Griffiths & Tenenbaum (2005) はこれら二つのモデルが本質的に異なる計算問題に対応していることを指摘しており、因果帰納における人間の推論がどちらの枠組みでなされるのか、またそれがどのように切り替わるのかは、現在も活発な議論の対象である。

Hattori et al. (2017) は、人間は状況に応じて直感的な推論 (A フレーム) と熟慮的な推論 (B フレーム) を切り替えるとする二重フレーム理論を提唱した。因

果帰納の文脈において、推論のどのような特性が A/B フレームの切り替えと対応しているかは議論の余地があるが、本研究では「因果の方向性を考慮するか否か」という推論形式の違いと対応していると仮定した。実際、因果強度モデルの枠組みで、否定の対称性（非対称／対称）と介入可能性の操作をすることで、人間の因果推論傾向の適合が無向なモデル DFH から有向なモデル  $\Delta P$  へ切り替わる現象が実証的に示されている (Hattori et al., 2017)。しかし、因果構造モデルにおいても同様の切り替えが生じるのかは検証されていない。

本研究では、情報の提示形式（オンライン／サマリー）と否定の対称性（非対称／対称）に着目し、これらの条件を操作することが人間の推論傾向と、因果の方向性を考慮する有向モデル (CS) ならびに方向性を考慮しない無向モデル (UCS) の適合にどのように作用するかを検証する。これにより、因果帰納の文脈における二重フレーム理論の整合性を、構造モデルの側面からも補完することを目指す。

### 2. 因果帰納

#### 2.1 単純因果帰納

単純因果帰納は、単一の原因候補事象 (C) が単一の結果事象 (E) を引き起こす度合いを評価する枠組みであり、表 1 に示す  $2 \times 2$  分割表に対応する事象の共起情報が用いられる。

表 1  $2 \times 2$  分割表

	Effect (E)	No Effect ( $\neg E$ )
Cause (C)	a	b
No Cause ( $\neg C$ )	c	d

#### 2.2 因果帰納の二重フレーム理論

因果帰納における人間の推論プロセスは、認知資源の効率的な利用と判断の正確性の間で適切なバランス

を取ることが必要とされる。Hattori et al. (2017) は、このバランスを説明する理論的枠組みとして「二重フレーム理論 (dual frame theory)」を提唱し、人間が状況に応じて異なる二つの認知的枠組み (フレーム) を使い分けて因果推論を行うことを示唆している。本理論は、以下の2つのフレームから構成される。

### 2.2.1 ヒューリスティック段階 (A フレーム)

A フレームでは、因果関係がある可能性のある事象の迅速な特定に重点が置かれる。この段階では、陽性事象 (*a* セル: 原因と結果がともに生起, *b* セル: 原因が生起し結果は非生起, *c* セル: 原因は非生起で結果が生起) に焦点を当て、*d* セル (原因も結果も非生起) の情報は軽視される (稀少性仮定)。

この段階において代表的なモデルは、Dual Factor Heuristic (DFH; Hattori & Oaksford, 2007) や proportion of Assumed-to-be Rare Instances (pARIs; Takahashi et al., 2010) である。

### 2.2.2 分析段階 (B フレーム)

B フレームでは、A フレームで特定された因果候補を対象に、介入や比較を通じて因果関係の存在と強度を厳密に評価する。全てのセル情報 (*a*, *b*, *c*, *d* セル) が考慮され、因果関係の規範的な評価が志向される。ここでは、陽性・陰性事象の対等な扱いや、原因・結果の非対称性の克服が重視される。

代表的なモデルとしては、 $\Delta P$  (Jenkins & Ward, 1965), Causal Power (Cheng, 1997), Causal Support (Griffiths & Tenenbaum, 2005) などがあり、これらは因果構造の識別および強度推定に基づく規範的判断を説明する。

フレームの切り替え ( $A \rightarrow B$ ) を誘発する要因としては、介入、対称的否定 (例: 「*X* を与える」と「*Y* を与える」)、判断への高いコミットメント (例: 経済的影響の大きい意思決定) などが指摘されており、これらが存在する場合、*d* セルも含む全情報を考慮する傾向が強まり、 $\Delta P$  モデルによる予測と整合する判断が増えることが実験的に示されている (Hattori et al., 2017)。

このように、因果帰納の二重フレーム理論は、人間が状況に応じてヒューリスティックな効率と分析的な正確性を適応的に切り替えることを説明する枠組みを提供する。

### 2.2.3 Causal Support (CS)

Causal Support (CS) は、特定の原因候補と結果の間に因果関係が存在するかを判断するためのベイズ的因果推論モデルである。本モデルは、観察データが「因果関係が存在する構造 (Graph 1)」と「因果関係が存在しない構造 (Graph 0)」のいずれをより支持するかを、

それぞれの尤度に基づいて比較するものである (図 2)。CS 値は両仮説の対数尤度比として定義され、値が大きいくほど因果関係の存在を強く支持するものと解釈される (Griffiths & Tenenbaum, 2005)。

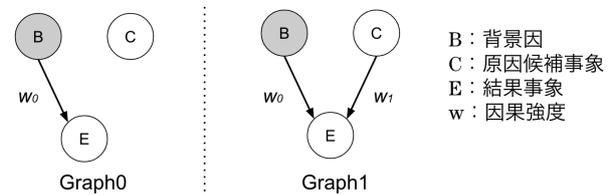


図 1 Causal Support で用いられる因果構造

### 2.2.4 Undirected Causal Support (UCS)

Undirected Causal Support (UCS) は、人間の因果帰納において、従来の有向グラフによる表現とは異なり、無向グラフ構造 (図 2) を用いて因果構造の学習を行うモデルである。CS と同様に、因果関係が存在しない仮説 (Graph 0') と存在する仮説 (Graph 1') の対数尤度比に基づいて、因果構造の支持度を評価する (樋口他, 2025)。

UCS は、因果の方向性を明示的に区別しない点や、「稀な事象に注目する」という稀少性仮定を取り入れている点において、DFH や pARIs と共通している。その意味で、UCS はこれらヒューリスティックモデルをグラフィカルモデルの枠組みへと拡張したものと位置づけることができる。

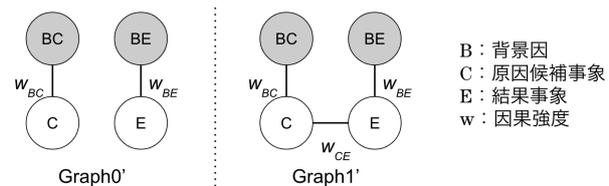


図 2 Undirected Causal Support で用いられる因果構造

## 3. 実験

### 3.1 実験概要

本研究では、二重フレーム理論の観点から、因果推論における推論形式の切り替えメカニズムを検討する。因果の方向性を考慮する合理的なモデル CS (B フレーム相当) と、因果の方向性を考慮しない稀少性ヒューリスティックを組み込んだ UCS (A フレーム相当) を用い、「情報の提示形式 (オンライン/サマリー)」と「否定の対称性 (非対称/対称)」からなる  $2 \times 2$  要因が、二つの推論形式 (有向/無向) の選択に影響を及ぼすのか、どのように影響を及ぼすかを検証する。



図3 オンライン形式の実験画面



図4 サマリー形式の実験画面

3.1.1 実験条件

対称的否定条件では、 $X$  と明示的に異なる原因  $Y$  を対比させ、非対称的否定条件では「 $\neg X$ 」のような曖昧な選択肢を背景として扱う（表2参照）。

オンライン形式（図3）では、 $a, b, c, d$  の各状況を逐次提示し、最後に因果関係の強度をスライダーで評価させた。視覚的な図版やキャラクターによる補助表示も含まれている。

一方、サマリー形式（図4）では、同一の内容を統計的に整理された表で集約して提示し、因果関係の強度を評価させた。

3.1.2 シナリオ

表2に示す12シナリオを作成した。参加者が状況を直感的に理解しやすくなるように、各シナリオには具体的な説明文と関連する画像を組み合わせて提示した。また、「あるサプリメント」や「サプリメントA/B」など、どちらが良い・悪いといった印象を与えない中立的な言葉や状況を用いることで、参加者が自分の判断に集中できるように設計した。

3.1.3 刺激設計

実験で使用した6種類の刺激とモデル値の対応関係については、表3に示す。なお、刺激に含まれる事象は、(i) UCSの値は一定でありながらCSの値は変化するようにするものであり、(ii) サマリー形式における二

表2 実験で用いたシナリオ一覧（非対称的な否定条件および対称的な否定条件）

シナリオ名	非対称的な否定	対称的な否定
サプリメントと頭痛	「あるサプリメント」を服用 vs 服用しない	「サプリメントA」 vs 「サプリメントB」
ソナーと海底資源発見	「ソナー」を使用 vs 使用しない	「A式ソナー」 vs 「B式ソナー」
土壌添加剤と花の開花	「ある土壌添加剤」を施す vs 施さない	「土壌添加剤A」 vs 「土壌添加剤B」
照明器具と植物の成長	「特定の照明器具」を設置 vs 設置しない	「照明A」 vs 「照明B」
新サービスと顧客満足度	「新サービス」を導入 vs 導入しない	「新サービスA」 vs 「新サービスB」
添加物とパンの膨らみ	「ある添加物」を加える vs 加えない	「A社の添加物」 vs 「B社の添加物」
ノイズ音と暗記テスト正答率	「微弱なノイズ音」を使用 vs 使用しない	「微弱なノイズ音A」 vs 「微弱なノイズ音B」
管理システムと病院運営効率	「新管理システム」を導入 vs 導入しない	「管理システムA」 vs 「管理システムB」
トレーニング器具と筋肥大	「新型トレーニング器具」を使用 vs 使用しない	「トレーニング器具A」 vs 「トレーニング器具B」
化学物質と燃費向上	「特定の化学物質」を加える vs 加えない	「化学物質A」 vs 「化学物質B」
色温度照明と睡眠の質	「特定の色温度の照明」を使用 vs 使用しない	「色温度A」 vs 「色温度B」
楽曲と授業への集中力	「ある楽曲」を聴く vs 聴かない	「楽曲A」 vs 「楽曲B」

つの群の事象の数の  $a + b$  と  $c + d$  が一致するように設計した。(i) は人間の推論傾向がCSとUCSのどちらと適合するかを検証するためであり、(ii) は  $a + b$  と  $c + d$  の合計の不一致による評価の難化を回避するためである。これらの要件をみだす刺激として、 $a$ セルを固定し、 $b$  および  $c$  の比率を操作したうえで、 $a + b$  と  $c + d$  の合計が一致するように  $d$ セルを決定した。

表3 刺激設計とモデルの予測値。表1に示した分割表を6通りに変化させた。

#	刺激設計				合計	モデルの予測値	
	$a$	$b$	$c$	$d$		CS	UCS
1	6	6	0	12	24	3.19	21.8
2	6	3	3	6	18	0.62	21.7
3	6	0	6	0	12	0.30	21.8
4	7	6	0	13	26	4.03	26.0
5	7	3	3	7	20	1.20	25.9
6	7	0	6	1	14	0.69	26.0

3.2 手続き

各参加者は、まず「否定の対称性」（非対称/対称）のいずれかの条件に割り当てられたうえで、オンライン形式とサマリー形式の2フェーズを経験し、各フェーズでは6つのシナリオがランダムな順序で提示された。また、提示形式と刺激の順序もランダムに決定され、2要因（オンライン/サマリー × 非対称/対称）の組み合わせが均等になるよう割り当てを制御した。各シナリオの提示後、参加者は、原因が結果に与える影響の大きさをスライダーで評価した（-100：強

い抑制, 100:強い促進)。

### 3.3 分析

順序効果を完全に排除するために, 今回の分析には最初に割り当てられた提示形式の回答値のみが使用された。この得られた結果をもとに, 4つの実験条件それぞれについて, シルエット法によって最適なクラスター数を特定し, k-means法を用いて参加者の評価傾向をクラスターリングによって分類した。そして, 全体及び各クラスターの平均回答値について因果構造モデルのうち Causal Support (CS) および Undirected Causal Support (UCS) のモデル値との相関係数を算出し比較した。

## 4. 結果と考察

参加者は CrowdWorks を通じて募集され, 同意を得た 200 名がオンライン実験に参加した。回答時間の極端値 (上位・下位 5%) および IMC (指示理解度チェック) に不通過の参加者を除外し, 最終的に 148 名のデータを解析対象とした。最適なクラスター数は 4 つの実験条件で一貫して  $k=2$  となった。条件×クラスターごとの平均評定値とモデルの相関係数を表 4 に示す。

結果から, オンライン×非対称の条件では, 参加者の評価が UCS 優勢のクラスター 1 ( $n=10, r=0.544$ ) と CS 優勢のクラスター 2 ( $n=29, r=0.616$ ) に分類された。一方, 否定を対称的にした場合 (オンライン×対称) や, 提示形式をサマリー形式に変更した場合 (サマリー×非対称) には, UCS 優勢のクラスターは見られず, すべてのクラスターで CS モデルとの相関係数が高かった。ただし, サマリー×非対称の条件では, CS と UCS の相関係数の差が  $|\Delta| < 0.30$  と小さく, 両モデルの差が曖昧であった。そして, サマリー×対称の条件では, UCS 優勢のクラスター 2 が再び観察された ( $n=6, r=0.654$ )。

これらの結果は, オンライン×非対称の条件において, 因果の方向性を考慮しない推論を行う参加者が一定数存在することを示している。一方, 提示形式または否定の対称性のいずれか一方を操作すると, ほぼ全員が因果の方向性を考慮した推論を行うようになることが明らかになった。さらに, 両方の要因を同時に操作した条件 (サマリー×対称) では, 無向的な推論を行う参加者が一定数維持されるという直観に反する結果が得られた。これは, 両要因の間に非加法的な相互作用が存在する可能性を示唆している。

Hattori et al. (2017) は, 人間の因果帰納推論が因果強度の推定に基づく場合に A/B フレームの切り替え

表 4 条件×クラスターごとの平均評定値とモデルの相関係数

提示	否定	クラスター	$n$	CS	UCS
オンライン	非対称	1	10	-0.387	0.544
		2	29	0.616	0.163
オンライン	対称	1	14	0.828	0.291
		2	24	0.832	0.181
サマリー	非対称	1	2	-0.287	-0.485
		2	37	0.721	0.433
サマリー	対称	1	26	0.786	0.147
		2	6	0.330	0.654

が生じることを実証した。これに対して本研究で得られた結果は, 因果構造の尤もらしさに基づいているという立場に立った場合にも, 同様のフレーム切り替えが起こることを示した。この結果は, 因果帰納における二重フレーム理論の妥当性と一般性をより強固に裏付けるものである。また, 有向的な推論・無向的な推論の切り替えは, 「情報の提示形式 (オンライン/サマリー)」によっても誘発されることが明らかとなった。

## 文献

- Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, 104(2):367–405.
- Griffiths, T. L. & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology*, 51(4):334–384.
- Hattori, I., Hattori, M., Over, D. E., Takahashi, T. & Bratgin, J. (2017). Dual frames for causal induction: the normative and the heuristic. *Thinking and Reasoning*, 23(3):292–317.
- Hattori, M. & Oaksford, M. (2007). Adaptive non-interventional heuristics for covariation detection in causal induction: Model comparison and rational analysis. *Cognitive Science*, 31(5):765–814.
- Jenkins, H. M. & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs: General and Applied*, 79(1):1–17.
- Takahashi, T., Kohno, Y. & Oyo, K. (2010). Causal induction heuristics as proportion of Assumed-to-be Rare Instances (pARIs). In *Proceedings of the 7th International Conference on Cognitive Science (ICCS2010)*, pages 361–362.
- 樋口 滉規・市野 弘人・高橋 達二 (2025). 記述性と合理性は両立できるか? 無向グラフィカルモデルに基づく因果帰納への構造的アプローチ。