

学位論文内容の要旨

報告番号	先端科学技術甲第182号	氏名	樋口 滉規
論文題目	単純因果帰納モデル pARIs の合理分析 および因果探索アルゴリズムへの応用		

我々を取り巻く環境は不確実性に満ちているが、我々は因果推論を通して、不確実な未来を予測し(あるいは過去を推測し)、時に未来を意図した方向に操作することができる。このように因果推論は人間の認知における本質的な構成要素の一つであるが、その認知過程には未解明な点が多い。人間の因果推論は時に誤りを含むが、一方でデータが少ない状況でも適切な因果関係を学習できるなどの長所がある。この学習過程をモデル化することは、人間を理解するという科学的な価値だけでなく、有用なシステムの開発に繋がる工学的な価値がある。

本研究は、人間が出来事の起こり方を観察することで因果関係を帰納的に推定する認知過程、因果帰納に関する研究である。著者は人間の因果帰納と高い適合を示す因果帰納モデル pARIsについて、過去に報告された複数の認知科学実験の結果を統計学的な整合性を保ちつつ統合するメタ分析によって人間の因果帰納との適合を定量的に分析した。また、合理分析という認知科学研究の方法論に基づき、網羅的なコンピュータ・シミュレーションを行うことで母集団パラメータの推定性能に関して定量的な分析を行った。また、一組の原因と結果の間の因果強度を取り扱うために用いられる因果帰納モデルの理論的な枠組みを拡張して、複数事象間の因果的な構造の推定アルゴリズムを開発し、提案したアルゴリズムについてコンピュータ・シミュレーションによって性能の検証を行った。論文は6章構成である。

第1章は序論であり、本研究で用いる3つの重要なフレームワークを導入した。1つ目が合理分析という認知科学研究の方法論である。この方法論は、人間を始めとする「行動主体の認知システムは、進化の過程において、その主体の行動を環境に対して最適化する」という前提に基づいて認知モデルを組み立てるものである。この方法論に基づき、人間の因果推論の仕組みもまた、我々を取り巻く環境に対して最適化されているという観点から、人間の因果推論の説明を行う。また、人間の因果推論過程を説明するモデルは、それが現実環境に適応したものであるかということによって評価される。2つ目が単純因果帰納という人間の因果推論研究のための枠組みである。これは2つの事象の共起の仕方に基づいて因果関係の強さを推定するという状況を想定した枠組みである。人間の因果推論に関する多くの認知科学研究がこの問題に取り組んでいる。3つ目が二段階仮説である。これは人間の因果推論を、出来事の観察に基づく推論の段階(ヒューリスティック段階)と、出来事を意図的に引き起こすなどして真の因果関係を特定する段階(介入段階)の二つ段階に分けて説明するものである。本研究は前者の段階のモデリングに焦点を当て、因果帰納モデルである「pARIsがどの程度人間と同様に振る舞うか」、また、「pARIsが現実環境においてどのような形で最適化されているか」を検証するものである。章の最後に、単純因果帰納における主要な4つのモデルの定義と性質を概説した。

第2章では、因果帰納モデルpARIsがどの程度人間と同様に振る舞うかをメタ分析に基づいて評価した。メタ分析とは、過去の文献・研究によって報告された複数の実験データを統計的に正しい方法で統合し、それらのデータに対して総合的な分析を行うための手法である。今回、人間の因果帰納を対象とした6つの文献で報告された認知実験の結果を固定効果モデルによって統合することでpARIsの記述性能を評価した。メタ分析の結果は、過去に提案された34個のパラメータを持たないモデルと8個のパラメータを持つモデルよりも、pARIsが人間の因果帰納をよく説明することを明らかにした。

第3章では、pARIsが現実環境に対してどのように最適化されているかをコンピュータ・シミュレーションによって分析した。pARIsを始めとする因果帰納モデルの役割は無数の事象の中から

因果関係がありそうな事象を絞り込むことである。このために有効な指針の一つには共分散関係の検出がある。事象間に因果関係がある場合は基本的に共分散関係にあるため、原因事象の絞り込みのために共分散を検出するという戦略は有効であると考えられる。したがって、因果帰納モデルの計算上の目的は母集団における共分散を検出することであると仮定し、その目的に際してpARIsがどの程度有効であるかをコンピュータ・シミュレーションにより分析した。その際、予め母集団パラメータから人工データを生成し、その人工データから母集団パラメータを推定する、パラメータ・リカバリーテストによってモデルの性能を評価した。人工データを生成するための母集団パラメータには、共変関係の規範的な指標であるピアソンの積率相関係数を用いた。結果は、pARIsが現実環境における制約を反映した幾つかの自然な仮定のもとで、その他の主要なモデルよりも母集団パラメータを高い精度で推定することを示し、したがって、pARIsは現実環境の構造に対して適応的な利点を有するモデルであることが明らかになった。

第4章では、前章では共変関係の検出であった因果帰納モデルの計算上の目的を、非独立性の検出であると仮定しなおして再度分析を行った。これは(非)独立性が共分散と同様に、因果関係がありそうな事象を絞り込むという目的に適した指標の一つであるためである。母集団の非独立性のパラメータを定義するために、確率論的な独立の定義を拡張した。著者は因果関係の推定の対象になるような出来事は基本的には稀な出来事であるという仮定を置き(なお、これを稀少性仮定と呼ぶ)、この仮定における(非)独立性の指標、DRを導出した。その後、DRを母集団パラメータとして設定した上で、再度コンピュータ・シミュレーションによる網羅的な分析を行った。コンピュータ・シミュレーションの結果は、pARIsが稀少性仮定の下での非独立性の指標であるDRの非常に優れた近似であること、また、母集団パラメータの優れた推定性能を有することを定量的に示した。

第5章では、単一事象間の因果関係の強さの推定という単純化された枠組みのためのモデルであったpARIsの理論的範囲を、複数事象間の因果的な構造の推定という、より複雑かつ現実的な因果探索の枠組みへと拡張した。ベイジアン・ネットワークは複数事象間の因果的な構造を有向非巡回グラフ(DAG)と条件付き確率表(CPT)で表すグラフィカルモデルであるが、観測データからに基づいてDAGのグラフ構造(=因果構造)を推定するために用いられる手法には多大な計算コストがかかるという課題がある。ベンチマークとしても用いられる代表的な構造推定アルゴリズムであるPCアルゴリズムは、ベイジアン・ネットワークにおいてグラフ構造と変数間の条件付き独立が密接な関係にあることを利用し、条件付き独立性の検定を用いてグラフ構造の推定を行うアルゴリズムである。ここで、第5章のコンピュータ・シミュレーションにおいてpARIsが優れた「非独立性」の検出モデルであることが示されたことに鑑み、pARIsの定義を確率論的に拡張した条件付きpARIsを導出し、これをPCアルゴリズムにおける条件付き独立性検定の代わりに用いることで、ベイジアン・ネットワークにおける構造推定アルゴリズムの提案を行った。また、提案したアルゴリズムに対して、コンピュータ・シミュレーションを用いて推定精度の評価を行った。コンピュータ・シミュレーションの結果は、提案したアルゴリズムは、PCアルゴリズムにおける条件付き独立性の検定のために標準的な手法であるカイ二乗検定を使用した場合と比べて、特に稀少性仮定下、あるいは観測データが少ない場合において優れた性能を有することを定量的に示した。

第6章では、上記した分析の結果を踏まえ、本論文の主張をまとめた。本研究は、メタ分析やコンピュータ・シミュレーションの結果を介して、人間の因果帰納における記述性能と現実環境と符合する特定の状況下での母集団推定性能の両面においてpARIsの有効性を示すものである。また、pARIsを条件付き独立性の検定に用いたベイジアン・ネットワークの構造学習アルゴリズムを開発することで、単純因果帰納という限定的な枠組みのためのモデルであったpARIsの理論的な枠組みを拡張した。提案したアルゴリズムについてコンピュータ・シミュレーションに基づく性能の検証を行った結果、単純因果帰納においてpARIsが示した性質と同様に、提案したアルゴリズムは稀少性仮定のもとで、また、サンプルサイズが小さい場合に特に有効であることが明らかになった。提案したアルゴリズムによる因果構造の推定が人間の構造学習とどの程度適合するかという点に関して、人間を対象とした実証実験に基づく分析が今後の課題である。