

# 経済因果チェーン検索システムの構築と応用

## Economic Causal Chain Search System and its Application

和泉 潔<sup>1</sup> 坂地 泰紀<sup>1</sup>

Kiyoshi Izumi<sup>1</sup>, Hiroki Sakaji<sup>1</sup>

<sup>1</sup> 東京大学大学院工学系研究科 システム創成学専攻

<sup>1</sup>Department of Systems Innovation, School of Engineering, the University of Tokyo

**Abstract:** In this research, we introduce a system for searching causal relationships of events related to economics and finance in a chain-like manner from a database extracted from economic text data. This system also introduces stock price analysis using this economic causal-chain search.

### はじめに

経済ニュース記事を読んでいると、株価の動きや商品の売り上げ、雇用や貿易など様々な経済事象に関する原因と結果の関係の記述がよく出てくる。例えば「将来の少子高齢化が〇〇を引き起こし、それに関連した〇〇に関する需要が増えることが予想される」とか「現在の株価下落は〇〇による市場参加者のリスク警戒心理の高まりを反映している」など、ある事象の経済的な波及効果やその事象の原因について述べている。

しかし経済現象のように人間の行動が関係する事象の因果関係を、数値データだけから統計的に解析することは難しい。なぜなら、人間が原因事象をどのように認識して、それに対してどのような行動を取るのかという行動ルールが因果関係の鍵となっているからである。時と場合によって因果関係が変化する。自然科学的現象のように数値データの統計分析で、客観的かつ普遍的な因果系列を取り出すことはほぼ不可能である。

そこで、本研究では人間が認識した因果関係が含まれていると思われる経済テキストデータを解析し、経済分野に関わる因果関係のデータベースを構築する。さらに、特定の事象を表すフレーズから派生する因果系列を検索する手法を開発する。本手法を用いて、ユーザが入力した語句に対して因果系列を表示し、ユーザが適切な系列に選択したり適切でない系列を削除したりできるシステムを実装した。本システムで作成された因果チェーン、経済的な波及効果と要因列挙が可能となる。

本稿では、因果チェーン検索システムと、本システムやアルゴリズムを用いた応用例を紹介する。

### 経済的因果関係の抽出とデータベース構築

人間が認知した因果関係について記述されたテキストデータを解析し、そこから因果関係を抽出する技術が必要となる。本システムでは、決算短信という上場企業が業績や財務状況を開示するために定期的に発行しているテキストから、手がかり表現を用いた手法[1]により因果関係を抽出した。

- 使用したテキストデータ：2012年10月から2018年5月に約2,300社が発行した約2万個の決算短信テキスト

- 抽出した因果関係：1,078,542個

抽出された因果関係は、決算短信の発行された日付、決算短信を発行した銘柄などと一緒にデータベースに保存される。

### 因果チェーンの構築アルゴリズム

上述の経済因果データベースから、特定の語句から関連する因果系列を構築する手法には、[2,3]を基に改良した手法を用いた。具体的には以下の4ステップからなる(図1)。

1. ユーザがフレーズを入力する。最初は終端ノードを入力テキストとする。
2. 終端ノードの表現と一定の類似性がある因果関係を経済因果データベースから選択し、ネットワーク追加候補とする。
3. Step 2で抽出したネットワーク追加候補と終端ノードの組み合わせについて因果関係ノード

間の類似度を計算する.因果関係ノード間の類似度が閾値 $\alpha$ 以上あるときにノードをネットワークに追加してネットワークを拡張する.

- Step 3 でネットワークに追加したノードを終端ノードとして Step 2, Step 3 を繰り返す.

因果チェーン検索には 2 種類の方向があり, 前向きの場合には入力テキストを最初の原因表現として, 原因-結果の順で波及効果を検索する(図 1a). 後ろ向きの因果チェーン検索の場合には, 入力テキストを最終の結果表現として, 結果-原因の順で潜在的な遠因を検索する(図 1b).

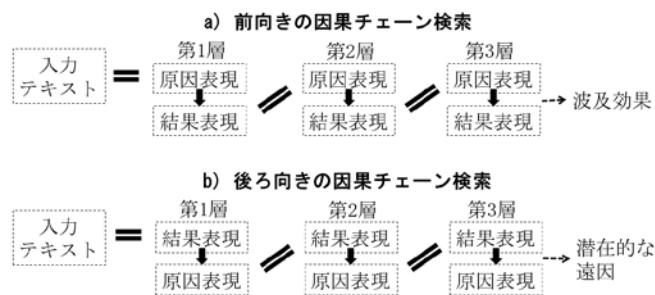


図 1. 前向きと後ろ向きの因果チェーン検索

## 経済因果チェーン検索システム

上述の因果チェーン構築アルゴリズムを基に, ユーザが提示された因果関係を編集できる機能を加えた経済因果チェーン検索システムを実装した<sup>1</sup>[4]. 以下に本システムの動作する様子を説明する.

最初に開始テキストをユーザが入力する(図 2). その際に, 原因から結果への波及効果の検索か結果から原因への要因検索かを指定する. さらに, 因果チェーンを構築する対象となる因果関係が発行された期間を限定することも可能である.

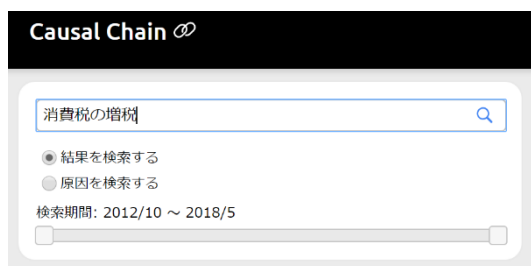


図 2. 開始テキストを入力するテキストボックス

テキストボックスの右にある検索ボタンをクリックすると, 開始テキストに連鎖する因果関係を表示

する(図 3). デフォルトでは関連性の高い順に 3 個の因果関係を表示する. より多くの因果関係を見たいければ, 「もっと見る」ボタンをクリックし因果関係ノードの表示を増やすことができる. 因果関係の各ノードに対して, ユーザが適切な因果でない判断した場合は, 各ノードの右上にある削除ボタンを押して, ノードを削除できる.



図 3. 連鎖する因果関係の表示

各ノードからさらに因果チェーンを伸ばしたい場合は, 各ノードの右側にある「>」ボタンをクリックすると, クリックされたノードを終端ノードとして関連する因果関係が追加される(図 4).

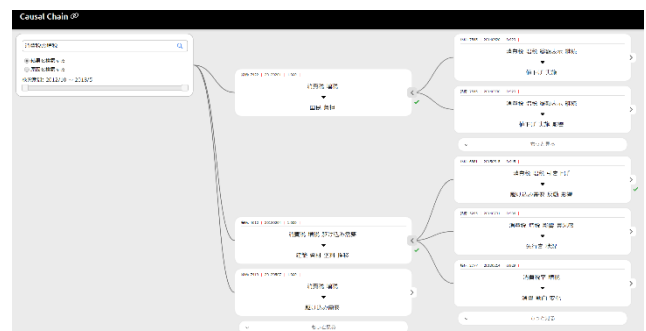


図 4. 因果チェーンの追加

上記の因果チェーンの構築作業を繰り返し, ユーザが必要とする因果チェーンを構築できたら, 構築した因果チェーンをファイルに保存することが出来る.

## 経済因果チェーンを用いた株価分析

経済因果チェーンの検索結果を用いて株価動向分析を行った研究事例を 2 つ紹介する.

### 因果関係を用いた銘柄間関連性指標

<sup>1</sup> 本システムは 2019 年 3 月より東京大学 和泉研究室のサイト (<http://socsim.t.u-tokyo.ac.jp>)にて公開予定である.

[6]は、本稿の経済因果チェーンの検索結果を持ちいて企業間の関連性の指標を計算した。ある企業 X の決算短信に出現した因果情報に後ろ向きの因果チェーン検索を行い、X の原因表現に類似する結果表現を有する因果情報を他の企業 Y の決算短信に出現する回数を X と Y の結びつきの強さとした。ある時点で企業 Y の株価に大きな変動があった場合に、市場参加者はその変動要因を類推し、変動要因に関連する企業 X の株式が売買されると仮定した。また Y から X への株価変動の伝播には、要因分析にかかるタイムラグが存在するとした。上記の仮定に基づき、TOPIX500 の構成銘柄に対して、先月の月次リターン(株価変動)が上位 20%の上昇した銘柄と因果関係で関連する銘柄を買った。逆に、先月に下位 20%の下落した銘柄に関連する銘柄を売るという売買を毎月行った。2012 年 12 月から 2019 年 1 月のデータで検証した結果、年率で 6.92%の正の利得が得られた。

### 因果関係を用いたニュースイベント分析

[7]は、特定のフレーズと関連する企業のリストを、因果チェーンの検索結果を用いて作成した。「小麦価格」という入力テキストに対して前向き因果チェーン検索の結果、図 A のような因果チェーンを得た。各層の因果情報を含む決算短信を発行した企業を調べて、各層で初めて出現した企業のリストを作成した。その結果、先ほどの「小麦 価格」から構築した因果チェーンでは、前向き方向と後ろ向き方向のどちらの検索結果でも、第 1 層では製粉会社や製パン業者など小麦価格に直接関連する企業が出現した。第 2 層、第 3 層には、小麦価格と間接的な関連を持つ企業が出現するようになった。例えば、前向き方向の検索の場合は、化学製品・調味料・食品製造機械メーカーなどが第 2 層と第 3 層で初出となった。後ろ向き方向の場合は、より広範な要因に関連する商社が出現した。このリストに含まれる企業に関して、2018 年 6 月以降で見出しに「小麦 価格」を含む大きなニュース記事が発行された時点の前後 20 営業日の価格変化を調べた。その結果、ニュース発行前に比べ発行後の価格変動が増大していた。さらに、より深い階層で出現した銘柄の株価の方がより大きな変動が長く持続した。

### まとめ

本研究では、経済・金融に関わる事象の因果関係を、経済テキストデータから抽出したデータベースから連鎖的に検索するシステムの紹介をした。本システムにより、ユーザが入力したイベントからの経

済的な波及効果を推論したり、ニュース記事からの因果チェーンをたどった関連銘柄を提示したりするなどのサービス応用のアイデアを紹介した。

### 謝辞

本システムを実装するにあたって、日名 洋平 氏と松田 健児 氏に多大な助力をいただいた。ここに深謝の意を表す。

### 参考文献

- [1] 坂地 泰紀, 酒井 浩之, 増山 茂: 決算短信 pdf から原因・結果表現の抽出, 電子情報通信学会論文誌 D, vol. 98, no.5, pp.811-822, 2015.
- [2] 西村 弘平, 坂地 泰紀, 和泉 潔: ベクトル表現を用いた因果関係連鎖の抽出, 人工知能学会 第 20 回金融情報学研究会, pp.50-53, 2018 年 3 月.
- [3] K. Nishumura, H. Sakaji, K. Izumi: Creation of Causal Relation Network using Semantic Similarity, 第 32 回人工知能学会全国大会, p. 1P104, 2018 年 6 月.
- [4] 坂地 泰紀, 和泉 潔, 松島 裕康: 金融テキストマイニングの基づいた投資家支援プラットフォームの開発, 人工知能学会 第 21 回金融情報学研究会, 2018 年 10 月.
- [5] 坂地 泰紀, 和泉 潔, 松島 裕康, 川瀬 和哉, 林 寛, 接触履歴を用いた地域景況インデックスの自動生成, 日本知能情報ファジィ学会誌 (to be appeared)
- [6] Nakagawa, K.; Sashida, S.; Sakaji, H.; Izumi, K. Economic Causal Chain and Predictable Stock Returns. In Proceedings of the 7th International Conference on Smart Computing and Artificial Intelligence (SCAI 2019), 655-660, 2019.
- [7] Kiyoshi Izumi, Shintaro Suda, and Hiroki Sakaji, Economic News Impact Analysis Using Causal-Chain Search from Textual Data, In Proceedings of KDF2020: The AAI-20 Workshop on Knowledge Discovery from Unstructured Data in Financial Services, 2020.

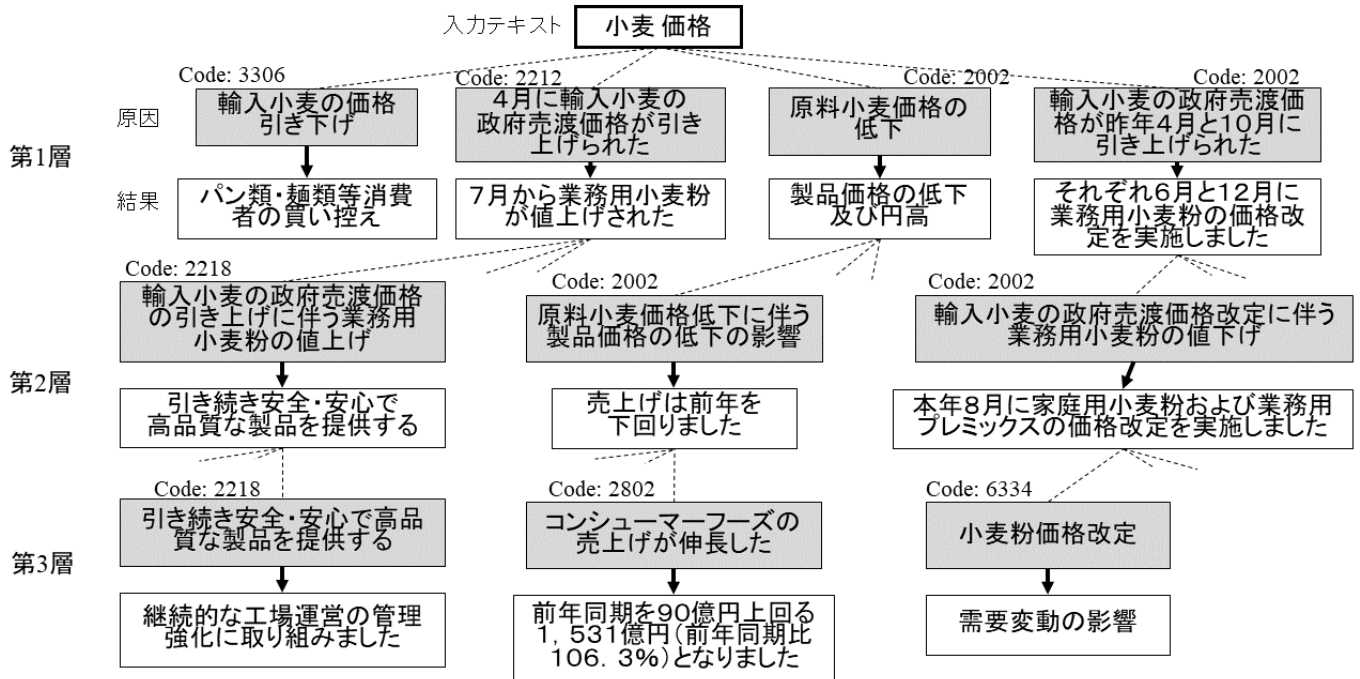


図 A 「小麦 価格」からの前向き因果チェーンの結果の一部