

明 細 書

発明の名称：知見抽出装置、知見更新装置、及びプログラム

技術分野

[0001] 本発明の実施形態は、知見抽出装置、知見更新装置、及びプログラムに関する。

背景技術

[0002] インターネットなどで提供される大量のデータを有効活用するために、ある概念に関する同義語や、概念間の関連を定義したシソーラスなどを用いて大量のデータからユーザに必要としている情報に関連する情報を選択して、もしくは選択した情報を組み合わせて知見情報を提供する技術がある。

[0003] 例えば、株取引においては、株変動のきっかけとなるような知見情報を迅速に入手して判断し、取引内容に反映させることが有効である。例えば、通常とは異なる時期に新型のインフルエンザが流行した場合、感染予防に使われるマスクを製造する企業や消毒薬を製造する企業の需要が増加することが容易に予想され、該当企業の株価が上昇する。株取引を行うユーザは、「インフルエンザの通常とは異なる時期の流行」の情報から、需要の増える銘柄を推測し、取引に反映することが有効である。

[0004] これを踏まえ、従来、証券情報と合わせてニュースを提供する技術や、ユーザの検索要求に応じて関連ニュースを抽出し、知見情報として提供する技術がある。

[0005] 上記のような技術として、例えば、株価チャートと関連ニュースを同一画面に表示する技術（第1の技術）や、注目している銘柄の関連銘柄（同業他社、取引先、メインバンク、資本関係あり、など予め設定したもの）のうち、注目している銘柄と同時期に同様の株価変動をした銘柄を関連銘柄として記憶する技術（第2の技術）がある。

[0006] また、ユーザの要求をトリガーとして記憶装置に保持されたニュース中の「～関連株」「～特需」などの手がかり語から注目銘柄検索のためのキーワ

ード抽出と銘柄関連付けを行い、ユーザに提示する技術や、一般的なシソーラス構築等の分野において、単語の係り受けを、複数発生する構文解析結果のうちの出現数に応じた確信度で管理・更新する技術がある（第3の技術）。また、文書群から企業名と共起する関連語を取得してテーブルに保持するとともに、その共起対が株価変動に影響する度合いを予測係数テーブルで保持し、関連語の評価値と、予測係数を、それぞれ管理・更新し、株価予測を行う技術（第4の技術）もある。

[0007] しかしながら、株価チャートとニュースを同一画面でユーザに提示する第1の技術では、ニュースに予め銘柄などの知見情報を抽出する対象の語句（以下、知見情報抽出対象語という）が関連付けられていなければならないという問題がある。

[0008] また、注目している銘柄と同時期に同様の株価変動をした銘柄を関連銘柄として記憶する第2の技術では、同様の変動をしなかった銘柄が関連付けられないという問題がある。また、注目している銘柄に同業などの明示的な関連のある銘柄を予め設定するため、銘柄をグループ化する範囲が限定されてしまう。

[0009] 第3の技術は、記憶装置に保持されたニュースからキーワード抽出と銘柄関連付けを実行するが、その活用方法や更新方法が開示されていない。

[0010] また、第3の技術は、単語の係り受けの確信度を評価し、以後に入手する新規コーパスで確信度を更新するが、情報源は単独（同種の情報）である。

[0011] 第4の技術では、企業名の関連語を取得・管理し、その株価変動への影響も管理・更新しているが、株価変動で更新する対象は予測係数のみである。

先行技術文献

特許文献

[0012] 特許文献1：特開2003-108785号公報

特許文献2：特開2003-162639号公報

特許文献3：特開平9-160915号公報

特許文献4：特開2011-141833号公報

発明の概要

発明が解決しようとする課題

[0013] 本発明が解決しようとする課題は、知見情報抽出対象に関連付けられていない状態で随時配信される電子文書から、当該知見抽出対象語に関する知見情報を抽出する知見抽出装置、知見更新装置、及びプログラムを提供することである。

課題を解決するための手段

[0014] 実施形態の知見抽出装置は、電子文書から知見情報を抽出する知見抽出装置であって、電子文書を受信する情報受信部と、知見情報を抽出する対象語と、知見情報を抽出する手掛かり語とに基づいて電子文書から概念を抽出し、抽出された当該概念と対象語とを関連付けた知見情報を作成する知見抽出部と、抽出された知見情報を記憶する記憶部と、知見情報が記憶された後、電子文書を記憶部内の知見情報に基づいて分析する情報分析部とを備える。

図面の簡単な説明

[0015] [図1] 第1の実施形態における知見抽出装置の構成例を示す図。
 [図2] 第1の実施形態における第1手掛かり語辞書の例を示す図。
 [図3] 第1の実施形態における知見抽出対象語辞書の例を示す図。
 [図4] 第1の実施形態における第2手掛かり語辞書の例を示す図。
 [図5] 第1の実施形態における電子文書の例を示す図。
 [図6] 第1の実施形態における電子文書の例を示す図。
 [図7] 第1の実施形態にしたがうユーザ端末装置の構成例を示す図。
 [図8] 第1の実施形態における知見抽出部の処理動作を詳細に示す図。
 [図9] 第1の実施形態における知見抽出部の処理動作の一例を説明するためのフローチャート。
 [図10] 第1の実施形態における処理動作の順序の概略を示す図。
 [図11] 第1の実施形態における分析対象の電子文書からの第1概念名であるグループ名、銘柄、株取引材料の抽出結果の一例を示す図。

[図12]第1の実施形態における知見情報の一例を示す図。

[図13]第1の実施形態における情報分析部の処理動作を詳細に示す図。

[図14]第1の実施形態における情報分析部の処理動作の一例を説明するため
のフローチャート。

[図15]第1の実施形態における情報分析部の処理動作の一例を説明するため
のフローチャート。

[図16]第2の実施形態における第2手掛かり語辞書の例を示す図。

[図17]第2の実施形態における電子文書の例を示す図。

[図18]第2の実施形態における知見情報の例を示す図。

[図19]第3の実施形態における知見抽出部の処理動作の一例を説明するため
のフローチャート。

[図20]第3の実施形態における知見抽出部の処理結果の一例を示す図。

[図21]第4の実施形態における知見抽出装置の構成例を示す図。

[図22]第4の実施形態における知見情報の一例を示す図。

[図23]第4の実施形態における知見評価更新装置の処理動作の一例を示す図。
。

[図24]第4の実施形態における知見評価更新装置の処理動作の一例を説明す
るためのフローチャート。

[図25]第4の実施形態における知見評価更新装置の処理動作の一例を説明す
るためのフローチャート。

[図26]第4の実施形態における知見評価更新装置の評価情報の一例を示す図。
。

[図27]第4の実施形態における知見評価更新装置の確信度更新ルールの一例
を示す図。

[図28]第4の実施形態における知見評価更新装置の確信度更新処理の一例を
示す図。

[図29]第4の実施形態における知見評価更新装置の確信度更新処理の一例を
示す図。

[図30]第4の実施形態における知見評価更新装置の確信度更新の効果の一例を示す図。

[図31]第5の実施形態における分析対象のニュース情報の例を示す図。

[図32]第5の実施形態における関連語情報の例を示す図。

[図33]第5の実施形態における知見抽出装置の構成例を示す図。

[図34]第6の実施形態における知見抽出装置の構成例を示すブロック図。

[図35]第6の実施形態における知見抽出装置の第1手掛かり語辞書の例を示す模式図。

[図36]第6の実施形態における知見抽出装置の知見抽出対象語辞書の例を示す模式図。

[図37]第6の実施形態における知見抽出装置の第2手掛かり語辞書の例を示す模式図。

[図38]第6の実施形態における知見抽出装置の電子文書の例を示す模式図。

[図39]第6の実施形態における知見抽出部の処理動作を詳細に示す模式図。

[図40]第6の実施形態における情報分析部の処理動作を詳細に示す模式図。

[図41]第6の実施形態における知見抽出部の処理結果の一例を示す模式図。

[図42]第6の実施形態における知見評価更新装置の処理結果の一例を示す模式図。

発明を実施するための形態

[0016] 以下、各実施形態の知見抽出装置について図面を用いて説明する。なお、以下の各装置は、装置毎に、ハードウェア構成、又はハードウェア資源とソフトウェアとの組合せ構成のいずれでも実施可能となっている。組合せ構成のソフトウェアとしては、予めネットワーク又は記憶媒体から対応する装置のコンピュータにインストールされ、対応する装置の機能を実現させるためのプログラムが用いられる。

[0017] (第1の実施形態)

図1は、第1の実施形態に係る知見抽出装置の構成例を示すブロック図である。本実施形態の知見抽出装置1は、株取引を行うユーザや株取引判断を

行うアルゴリズムトレードエンジンなどのシステム（複合イベント処理エンジン、CEPなどを組み込んだものを含む）への証券市場立会時間中の情報提供に使用可能なものであり、例えばインターネット等で配信されたニュース情報などの電子文書から株取引に関する知見情報を抽出する。抽出結果の、知見情報は情報提供などに使用可能となっている。

[0018] 具体的には、この知見抽出装置１は、ニュース情報のうちの経済ニュースを受信した際に、例えば、キーワード「インフルエンザ」に関連する銘柄グルーピングの知識や、「インフルエンザ」関連銘柄で株価変動の材料となる「集団感染」「鳥インフルエンザ」の知識を獲得し、これらの知識を知見情報として保持するものである。また、知見抽出装置１は、経済ニュースを含むインターネット配信ニュース、企業情報を受信した際は、保持した知見情報を参照し、ニュース情報に「インフルエンザ」などのグループ、該当グループに所属する銘柄、該当グループの株価変動材料の有無の情報を付与し、アルゴリズムトレードエンジンなどのシステムもしくは株取引を行うユーザのユーザ端末装置に提供可能となっている。すなわち、知見抽出装置１は、銘柄などの知見抽出対象語のグルーピングをするための情報である概念（第１概念）や、第１概念に関連づける情報（第２概念）を含む知見情報を抽出する。

[0019] 例えば、株取引に関する知見情報を抽出する場合、株の銘柄が所属する業界（例：金融、金属、広告など）が第１概念となるような手掛かり語（第１手掛かり語）を、知見抽出装置１に記憶させる。第１手掛かり語については後述する。第２概念は、第１概念と異なる概念であり、後述する第２手掛かり語に基づいて抽出される。

[0020] また、知見抽出装置１は、知見抽出システム１と呼んでもよい。

[0021] 具体的には知見抽出装置１は、装置全体の制御を司る制御部１０、記憶部１１、情報受信部１２、情報選択部１３、知見抽出部１４、情報分析部１５および入出力インタフェース１６を備え、各部１１～１６がバス１７を介して接続される。

[0022] 記憶部 11 は、例えばハードディスクドライブや不揮発性メモリ装置などのハードウェアで構成される。記憶部 11 は制御部 10 による実行対象のプログラムを記憶するのに加え、知見抽出部 14 で抽出した知見情報、受信し分析を行った電子文書などを保持する。また、記憶部 11 は、後述する第 1 手掛かり語辞書 111、知見抽出対象語辞書 112、第 2 手掛かり語辞書 113 及び処理途中のデータ等も保持する。

[0023] 第 1 手掛かり語辞書 111 は、図 2 に一例を示すように、第 1 手掛かり語を識別するコードと、知見情報抽出対象をグルーピングする第 1 概念の手掛かりとなる第 1 手掛かり語とが関連付けられて記述されている。すなわち、本実施形態の電子文書はあらかじめ記憶された第 1 手掛かり語に基づいて第 1 概念名（グループ名）をつけられる。本実施形態の第 1 手掛かり語辞書はグループ定義語辞書 111 といえる。

[0024] 知見抽出対象語辞書 112 は、図 3 に一例を示すように、本実施形態において電子文書から知見情報を抽出する対象語（知見抽出対象語）である銘柄を識別する証券コードと、銘柄の企業名と、銘柄の業種の大分類を示す東証業種 1 と、東証業種 1 内の小分類を示す東証業種 2 と、銘柄の企業名を表す企業名表現とが関連付けられて記述されている。企業名表現としては、企業名の正称、略称、通称、俗称、愛称などの様々な名称が使用可能となっている。なお、企業名表現は、名称に限らず、各企業の証券コードを用いてもよい。すなわち、本実施形態では知見抽出対象語辞書を銘柄辞書 112 といえる。

[0025] 第 2 手掛かり語辞書 113 は、図 4 に一例を示すように、第 2 手掛かり語（材料表現手掛かり語）を識別するコードと、後述する知見抽出判定語の表現に対する手掛かり語とが関連付けられて記述されている。すなわち、本実施形態では第 2 手掛かり語辞書 113 を材料表現手掛かり語辞書 113 といえる。

[0026] 情報受信部 12 は、インターネットで配信されるニュース情報を受信する。図 5 及び図 6 はインターネットで配信されるニュース情報の例である。ニ

ユース情報の受信は、配信業者と契約を結び配信を受ける形態でも、一般に検索ロボットもしくはクローラと呼ばれるインターネット検索プログラムにより新規配信されたニュース情報を発見し入手する形態でもよい。本実施形態では、インターネットで配信されるニュース情報の構成要素のうち、図5及び図6に示す「タイムスタンプ」「ニュースソース」「ジャンル」「ニュース見出し」「ニュース本文」の5項目と、情報受信部12が付与する「ニュースID」を利用する。「タイムスタンプ」は、配信されたニュース情報に付与された配信日時情報である。「ニュースソース」は、その情報を提供した事業者の名称である。「ジャンル」は、ニュース情報に付与された「経済」「社会」「スポーツ」などのジャンル（分野）を示すタグ情報である。「ジャンル」は、ニュース受信時に付与されていなくてもよい。「ニュース見出し」は配信されたニュース情報の見出し部分の日本語列、「ニュース本文」は配信されたニュース情報の見出しを除く本文部分の日本語列である。

[0027] 情報選択部13は、情報受信部12が受信したニュース情報から、ここでは経済ニュース（経済用語を含むニュース情報）を選択する。経済ニュースであることの判断は、例えば、受信したニュース情報に「ジャンル」が付与されている場合は、「ジャンル」が「経済」「市況」などの経済とその類義語のうちのいずれかの用語を含む表現であることにより判断する。受信したニュース情報に「ジャンル」が付与されていない場合は、例えば、ニュース見出しもしくはニュース本文に「東証」「大証」「日経平均」「株価」「株式市場」「高値」「終値」「出来高」などの経済用語、特に株取引に関する用語のいずれかを含むことにより判断する。すなわち、情報選択部13は、知見抽出対象語に関連の高い電子文書を選択する。

[0028] 知見抽出部14は、情報入力部13によって選択されたニュース情報から知見情報を抽出するものであり、例えば、後述する知見抽出判定語チェック部141、第1手掛かり語チェック部142、第1概念抽出部143、知見抽出対象語抽出部144、第2概念抽出部145、組み合わせ作成部146及び第1概念チェック部147により、記憶部11内の各辞書111～11

3を用い、知見情報抽出対象の経済ニュースから、知見抽出対象である銘柄をグループ分けする第1概念名、グループに所属する銘柄、株価変動の要因となる材料表現を抽出し、銘柄グルーピングや材料（イベント）の知見情報として記憶部11に書込む。ここで、知見情報としては、例えば、第1概念名、銘柄及び株価変動材料を含む情報、又は第1概念名及び銘柄からなる第1情報と株価変動材料からなる第2情報とを含む情報が使用可能となっているが、これらの組み合わせには限定されない。なお、知見抽出部14は情報選択部13によって選択された情報ではなく記憶部11に記憶された電子文書から知見情報を抽出することも可能である。

[0029] なお、知見抽出部14は、これら各部141～147及び各辞書111～113を用いる方法に限らず、例えば、構文解析、意味解析、文脈解析を用いて銘柄をグループ分けする方法により、知見情報を抽出してもよい。後者の方法では、例えば「インフルエンザの流行を受けて、XX社がランクイン」のように、図2に示したグループ定義語を含まないニュース情報からも知見情報を抽出可能となっている。

[0030] 情報分析部15は、知見情報が記憶部11に記憶された後、情報受信部12に（新規に）受信された全てのニュース情報を記憶部11内の知見情報に基づいて分析するものであり、例えば、後述する知見抽出対象語・第1概念・第2概念抽出部151、電子文書割付部152及び知見抽出対象別判定部153を備えている。また、情報分析部15は、新規に受信したニュース情報にグルーピングや材料（イベント）の知見情報に該当する情報が存在する場合は、当該ニュース情報に当該知見情報を付与してもよい。情報分析部15はさらに、新規に受信したニュース情報を含めて一定時間に処理したニュース情報をグループや銘柄にわけ、それぞれのグループや銘柄で統計処理を行い、統計情報を付与してもよい。

[0031] 一連の処理終了後、情報分析部15の出力が、ニュース情報の分析結果として入出力インタフェース16によりユーザ端末装置に提示される。

[0032] 入出力インタフェース16は図示しない外部記憶装置とケーブルを介して

接続可能であり、この外部記憶装置との間で記憶部 11 に記憶される知見情報や分析結果保存データベース（図示せず）にデータを入出力する。

[0033] 図 7 は、以上のような知見抽出装置 1 に接続されるユーザ端末装置の構成例を示すブロック図である。ユーザ端末装置 2 は、知見抽出装置 1 を利用するユーザが取り扱う端末装置である。ユーザ端末装置 2 は、装置全体の処理を司る端末制御部 21、端末記憶部 22、表示部 23、入力部 24、通信インタフェース 25 を備え、それぞれがバス 26 を介して相互に接続される。

[0034] 端末記憶部 22 は、例えばハードディスクドライブや不揮発性メモリ装置などのハードウェアで構成された記憶装置である。端末記憶部 22 は、制御用のプログラムを記憶する。また、端末記憶部 22 は、端末制御部 21 による各種処理のワークメモリとしても機能する。

[0035] 表示部 23 は、例えば液晶ディスプレイであり、情報分析部 15 の出力をユーザに向けて出力する。例えば、新規に受信したニュース情報の配信時刻（タイムスタンプ）、ニュースソース、ニュース情報で話題とされた企業名・第 1 概念名、ニュース情報に含まれる株取引材料（自己株取得、経営悪化、インフルエンザの集団感染、など）を表形式でユーザに提示する。

[0036] 入力部 24 は例えばキーボードやマウスであり、知見抽出、情報分析およびシステム設定にかかる操作を受け付ける。

[0037] 通信インタフェース 25 は、ケーブルを介して知見抽出装置 1 と接続され、ユーザ端末装置 2 においてユーザが指定した条件や知見抽出装置 1 の処理結果などのやり取りを行う。また、通信インタフェース 25 は外部記憶装置とケーブルを介して接続可能であり、この外部記憶装置との間で記憶部 11 に記憶する銘柄グルーピングや材料（イベント）の知見情報、分析結果を入出力することもできる。

[0038] 次に、以上のように構成された知見抽出装置 1 の動作について説明する。図 8 は、知見抽出部 14 の処理動作の詳細を示す模式図、図 9 は知見抽出部の処理動作の一例を説明するためのフローチャートである。

[0039] なお、以下の説明では、知見抽出部 14 の処理を、ニュース情報を受信す

る都度としているが、この処理は夜間などの情報分析処理を行わない時間帯に、受信済みのニュース情報に対してまとめて実施してもよい。

[0040] 情報受信部 12 が、ニュース配信サイトからの配信を受けるか、ニュース配信サイトにアクセスすることにより、新規のニュース情報を受信する。例えば、図 5 及び図 6 のうち、図 5 に示すニュース ID “1” に該当するニュース情報を受信する。

[0041] 情報選択部 13 は、情報受信部 12 が受信した電子文書であるニュース情報から、知見情報を抽出する電子文書として経済ニュースを選択する（ステップ S1）。経済ニュースであることの判断は、例えば、受信したニュース情報に「ジャンル」が付与されている場合は、「ジャンル」が「経済」「市況」などの経済とその類義語のうちのいずれかを含む表現であることにより判断する。

[0042] 受信したニュース情報に「ジャンル」が付与されていない場合は、例えば、ニュース見出しもしくはニュース本文に「東証」「大証」「日経平均」「株価」「株式市場」「高値」「終値」「出来高」などの経済用語、特に株取引に関する用語のいずれかを含むことにより判断する。選択されたニュース情報は、情報選択部 13 から知見抽出部 14 に送出される。

[0043] 知見抽出部 14 においては、知見抽出判定語チェック部 141 が、情報選択部 13 に選択されたニュース情報が知見情報を抽出する可能性のある電子文書かどうかを判定する知見抽出判定語を含むか否かを判定する（ステップ S2）。

[0044] 本実施形態では知見抽出判定語は株価変動を評価する株価変動評価語であるとする。すなわち、本実施形態では知見抽出判定語チェック部 141 を株価変動評価語チェック部 141 といえる。

[0045] 例えば、知見抽出判定語チェック部 141 は、受信したニュース情報のニュース見出しもしくはニュース本文に、「高値更新」「強い動き」「反発」「急落」「続伸」「ランクイン」などの株価変動評価語が含まれるか否かを判定する。

- [0046] ニュース情報に株価変動評価語が含まれる場合、知見抽出部 14 は、当該ニュース情報が知見情報を含むニュース情報であると判定して、ニュース情報を第 1 手掛かり語チェック部 142 に送信する。
- [0047] なお、「株価変動評価語」は、株価の変動を評価する内容であれば、「株価変動表現語」、「株価変動語」又は「株価変動定義語」といった他の用語に言い換えてもよい。
- [0048] 第 1 手掛かり語チェック部 142 は、ステップ S2 の判定の結果、ニュース情報が株価変動評価語を含むとき、当該ニュース情報が銘柄の第 1 概念名を定義する第 1 手掛かり語（グループ定義語）を含むか否かを判定する（ステップ S3）。すなわち、本実施形態の第 1 手掛かり語チェック部 142 はグループ定義語チェック部 142 といえる。
- [0049] 例えば、第 1 手掛かり語チェック部 142 は、第 1 手掛かり語辞書 111 を参照し、受信したニュース情報のニュース見出しもしくはニュース本文にグループ定義語の少なくとも 1 つが含まれるか否かを判定する。例えば図 5 に示すニュース ID “1” のニュースは、グループ定義語「関連銘柄」を含む。
- [0050] 受信したニュース情報にグループ定義語が含まれない場合は第 1 概念抽出部 143 によるステップ S4 の処理を省略する。
- [0051] 受信したニュース情報にグループ定義語が含まれる場合は、第 1 手掛かり語チェック部 142 は、ニュース情報を第 1 概念抽出部 143 に送信する。
- [0052] 第 1 概念抽出部 143 は、ステップ S3 の判定の結果、ニュース情報がグループ定義語を含むとき、当該グループ定義語に基づいて当該ニュース情報から第 1 概念名を抽出する（ステップ S4）。
- [0053] 例えば、第 1 概念抽出部 143 は、グループ定義語を含むニュース情報を受信すると、グループ定義語直前の名詞を第 1 概念名として抽出する。受信したニュース情報が例えば図 5 に示すニュース ID “1” の場合は、グループ定義語「関連銘柄」の直前の名詞「インフルエンザ」を第 1 概念名として抽出する。

- [0054] 第1概念名は名詞としたが、「インフルエンザ」のような単語ではなくて、「新型インフルエンザ」「季節性インフルエンザ」のような複合名詞も抽出対象としてよい。また、グループ定義語を含むが直前が名詞でない場合は、第1概念名を含まないと判断する。
- [0055] 知見抽出対象語抽出部144は、知見抽出判定語チェック部141によるステップS2の判定の結果、ニュース情報が株価変動評価語を含むとき、当該ニュース情報から企業名及び証券コードを含む銘柄を抽出する。
- [0056] 具体的には、知見抽出対象語抽出部144は、例えば、ステップS3又はS4の後、第1概念抽出部143の処理を省略して送信されたニュース情報に対して、知見抽出対象語辞書112を参照し、ニュース見出しもしくはニュース本文に知見抽出対象名表現が含まれるか否かを判定し（ステップS5）、知見抽出対象名表現が含まれる場合に当該知見抽出対象名表現に対応する知見抽出対象語である銘柄を抽出する（ステップS6）。すなわち、本実施形態の知見抽出対象語抽出部144は銘柄抽出部144といえる。
- [0057] ニュース見出しまたはニュース本文に、知見抽出対象語辞書112の「知見抽出対象名表現」列の表現がある場合は、その行の「証券コード」「企業名」に記された知見抽出対象語である銘柄が含まれると判断する。
- [0058] 知見抽出対象語辞書112は、1つの証券コードに対して複数行の定義を含んでよい。また、1件のニュース情報から複数の知見抽出対象名表現を抽出してよい。
- [0059] 例えば図5のニュースID“1”のニュース情報の場合、知見抽出対象名表現「RRボウ」「yyボウHD」「ホヘト薬品」が含まれるため、銘柄「証券コード 1031 企業名 RRボウ（株）」「証券コード 1033 企業名 yyボウホールディングス（株）」「証券コード 1041 企業名 ホヘト薬品（株）」が抽出される。
- [0060] 知見抽出対象語辞書112では知見抽出対象名表現を企業名称もしくは企業名の略称としているが、各企業の証券コードも知見抽出対象名表現の一例としてもよい。本実施形態の知見抽出対象語辞書112における知見抽出対

象名表現は企業名表現といえる。

- [0061] なお、知見抽出対象語抽出部 144 は、図 10 に示すように、ステップ S 2 の結果、株価変動評価語を含む場合に銘柄を抽出すればよいので、ステップ S 3, S 4 の後に限らず、ステップ S 2 の後からステップ S 9 の前の間の任意の時点で銘柄を抽出可能である。
- [0062] 第 2 概念抽出部 145 は、知見抽出判定語チェック部 141 による判定の結果、ニュース情報が知見抽出判定語である株価変動評価語を含むとき、当該ニュース情報から第 2 概念として株価変動材料を抽出する。具体的には、第 2 概念抽出部 145 は、株価変動評価語を含むニュース情報のニュース見出しもしくはニュース本文に、第 2 手掛かり語辞書 113 内の第 2 手掛かり語の 1 つ以上が含まれるか否かを判定する（ステップ S 7）。
- [0063] ステップ S 7 およびステップ S 8 の処理は、図 9 のフローチャートではステップ S 6 の後としているが、ステップ S 3 の処理の前、もしくはステップ S 3 からステップ S 6 の処理と並列で実施してもよい。
- [0064] 第 2 概念抽出部 145 は、受信したニュース情報のニュース見出しもしくはニュース本文に、第 2 手掛かり語辞書 113 内の第 2 手掛かり語が含まれる場合、その第 2 手掛かり語と係り受けしている名詞句を第 2 概念名として抽出する（ステップ S 8）。本実施形態では第 2 概念名を株価変動材料表現といえる。すなわち、本実施形態では第 2 概念抽出部 145 は材料抽出部 145 といえる。
- [0065] 例えば、図 5 に示すニュース ID “1” のニュース本文では、第 2 文「〇県が×市内の学生寮において、新型インフルエンザの集団感染が発生したと発表したことが手掛かり材料となっているようだ。」に第 2 手掛かり語「手掛かり材料」が含まれる。
- [0066] 第 2 手掛かり語「手掛かり材料」と係り受けしている名詞句は「新型インフルエンザの集団感染が発生したと発表したこと」である。この名詞句「新型インフルエンザの集団感染が発生したと発表したこと」が株価変動材料（イベント）表現として抽出される。

- [0067] 株価変動材料（イベント）表現は、抽出した名詞句を分析し、「新型インフルエンザの集団感染が発生した」もしくは「新型インフルエンザの集団感染」を抽出対象としてもよい。
- [0068] また、ステップS 6、ステップS 7で抽出する第1概念名「インフルエンザ」を含む「新型インフルエンザ」とその付属語を消去し「集団感染」を抽出対象としてもよい。
- [0069] さらに、該当のニュース本文が「〇県が×市内の学生寮において、新型インフルエンザの集団感染が発生したと発表した。これが手掛かり材料となっているようだ。」のように2つの文にわかれている場合、手掛かり語「手掛かり材料」と係り受けしている代名詞「これ」を抽出した後で、文脈解析により「これ」を前文の「〇県が×市内の学生寮において、新型インフルエンザの集団感染が発生したと発表した」に置き換え、置き換えた後の「〇県が×市内の学生寮において、新型インフルエンザの集団感染が発生したと発表した」もしくはその分析結果から「新型インフルエンザの集団感染が発生した」、「新型インフルエンザの集団感染」、もしくは「集団感染」を抽出対象としてもよい。
- [0070] ステップS 3からステップS 8の処理の終了後、組み合わせ作成部146は、処理中のニュース情報から抽出された第1概念名、知見抽出対象語である銘柄及び株価変動材料（第2概念名もしくはイベント）を組み合わせで知見情報を作成し、当該知見情報を記憶部に書き込む（ステップS 9）。
- [0071] ニュース情報から抽出された第1概念名が1つである場合は、その第1概念名に、抽出された全ての銘柄（知見抽出対象語）、及び抽出された全ての株価変動材料（第2概念名）を割り付ける。すなわち、1つの第1概念名に対して、知見抽出対象語である銘柄と第2概念である株価変動材料はそれぞれ1つまたは複数割り付けられることを許容する。
- [0072] 分析対象のニュース情報が図6のニュースID“3”の場合は、図11に示すようにニュース本文から複数の6件の第1概念名、8件の銘柄、3件の株価変動材料が抽出される。

[0073] このように抽出された第1概念名が複数である場合、文脈解析などの高度な自然言語処理を用いて、第1概念名、第1概念名に係り受けしている銘柄、第1概念名もしくはその第1概念名に係り受けしている銘柄に係り受けしている株価変動材料を組み合わせてもよい。

[0074] また、簡易的には、ニュース本文を文に区切り、文ごとに第1概念名、銘柄、株価変動材料を組み合わせるとしてもよい。さらに、ニュース本文を1文ずつ取り出し、第1概念名と銘柄が揃った時にそこまでに抽出された第1概念名、銘柄、株価変動材料を組み合わせで出力し、次の文から新たに第1概念名、銘柄、株価変動材料を組み合わせるとしてもよい。知見情報は第1概念名に対して銘柄と株価変動材料が割り付けられるとしているが、銘柄もしくは株価変動材料がない組み合わせが作成されてもよい。

[0075] このような処理により、本実施形態では、第1概念名、銘柄（企業名表現）、株価変動材料のセットであるされた知見情報として以下の6セットが抽出される。

[0076] (1) 第1概念名：石油 銘柄：○×△石〈1001〉、石油VV〈1002〉、

材料：10年4～12月期で連結経常益2000億円強が報じられた。

[0077] (2) 第1概念名：商社 銘柄：HH商〈1080〉。

[0078] (3) 第1概念名：非鉄金属 銘柄：SS鋳〈1050〉。

[0079] (4) 第1概念名：繊維 銘柄：センイJJ〈1030〉。

[0080] (5) 第1概念名：銀行 銘柄：HH銀〈1082〉。

[0081] (6) 第1概念名：ハイテク 銘柄：ハイテクBB〈1060〉、QQメモリ〈1063〉、

材料：D証（D証券）が投資判断を引き上げ。

[0082] 組み合わせは、第1概念名、銘柄、株価変動材料をセットにするとしたが、組み合わせは第1概念名と銘柄のみとし、株価変動材料は単独で知見情報として保持するとしてもよい。または、株価変動材料は第1概念名ではなく銘柄

柄とセットにして知見情報として保持するとしてもよい。

- [0083] 組み合わせ作成部 146 の処理後、図 12 に一例を示す如き、抽出された知見情報は、記憶部 11 に書き込まれる（ステップ S9）。なお、図 12 に示すように、本実施形態における第 1 概念の例として「インフルエンザ」、「石油」、「商社」、「非鉄金属」、「繊維」、「金属」、「ハイテク」などがある。
- [0084] ステップ S2 で株価変動評価語がないと判定された経済ニュースは、第 1 概念チェック部 147 により、ニュース見出しもしくはニュース本文に、記憶部 11 に書き込まれた知見情報の第 1 概念名（既存第 1 概念名）を含むか否かを判定される（ステップ S11）。本実施形態では第 1 概念名はグループ名であるため、第 1 概念チェック部 147 はグループ名チェック部といえる。
- [0085] 既存第 1 概念名を含む場合は、知見抽出対象語抽出部 144 が知見抽出対象語辞書 112 を用いて、ニュース見出しとニュース本文に知見抽出対象名表現が含まれるか否かを判定する（ステップ S12）。
- [0086] 知見抽出対象名表現が含まれる場合、例えば「参入」「新事業」などの表現がある場合は、組み合わせ作成部 146 が抽出された知見抽出対象名表現に対応する銘柄を既存グループに追加する処理のための組み合わせを作成する。
- [0087] 「撤退」などの表現がある場合は、組み合わせ作成部 146 が、既存第 1 概念に対して抽出された知見抽出対象名表現に対応する銘柄を取り除く処理のための組み合わせを作成する（ステップ S13）。
- [0088] ステップ S13 で既存第 1 概念と銘柄の追加もしくは削除処理の組み合わせが作成された場合は、その処理を記憶部 11 に対して実行する（ステップ S9）。
- [0089] 次に、情報分析部 15 の動作を図 13、図 14 及び図 15 によって説明する。
- [0090] 始めに、情報受信部 12 は、ニュース配信サイトからの配信を受けるか、

ニュース配信サイトにアクセスすることにより、新規のニュース情報を受信する。例えば、図5のニュースID2に該当するニュース情報を受信する（ステップS21）。受信したニュース情報は知見抽出対象語・第1概念・第2概念抽出部151に送信される。

[0091] 知見抽出対象語・第1概念・第2概念抽出部151は、まず、記憶部11に記憶された知見抽出対象語辞書112を参照し、受信したニュース情報のニュース見出しもしくはニュース本文に知見抽出対象名表現があるか否かを判定する（ステップS22）。

[0092] 知見抽出対象名表現がある場合、知見抽出対象語・第1概念・第2概念抽出部151は、抽出された知見抽出対象名表現に対応する企業名・証券コードを抽出する。また、知見抽出対象語・第1概念・第2概念抽出部151は、図12に示す知見情報を参照し、抽出した企業名・証券コードが割り付けられた第1概念名を全て抽出する（ステップS23）。知見抽出対象語・第1概念・第2概念抽出部151及び電子文書割付部（ニュース割付部）152は、ニュース見出しとニュース本文から抽出された企業名ごとに、その企業が属する各グループに対して、ステップS24・ステップS25の処理を行う。

[0093] 始めに、知見抽出対象語・第1概念・第2概念抽出部151は、図12に示す知見情報を参照し、その時点で処理中の企業名・その企業が属する第1概念名に割り付けられた第2概念（株価変動材料）を抽出し、ニュース見出し・ニュース本文中に該当の株価変動材料（表現）があるか否かを判定する（ステップS24）。

[0094] 該当の株価変動材料（表現）がある場合は、処理中の受信ニュース情報にそのグループに関する情報があると判断し、電子文書割付部152が、処理中のニュース情報を、図12に示す知見情報でその第1概念名に対応付けられた企業に割り付ける（ステップS25）。

[0095] 知見抽出対象語・第1概念・第2概念抽出部151は、ステップS22からステップS25の処理の後、もしくは、前、または並列で、受信したニュ

ース情報のニュース見出しもしくはニュース本文に、図 1 2 に示す知見情報内の第 1 概念名が含まれるか否かを判定する（ステップ S 2 6）。

[0096] 例えば、図 6 に示すニュース ID “2” のニュース情報では、本文に第 1 概念名「インフルエンザ」が含まれる。

[0097] 第 1 概念名が含まれる場合、知見抽出対象語・第 1 概念・第 2 概念抽出部 1 5 1 は、当該第 1 概念名を本文又は知見情報から抽出して電子文書割付部 1 5 2 に送出する。

[0098] 第 1 概念名が抽出された場合、電子文書割付部 1 5 2 は、図 1 2 に示す知見情報でその第 1 概念名に対応付けられた知見抽出対象語である銘柄を取り出し、処理中の電子文書を該当知見抽出対象語全てに割り付ける（ステップ S 2 7）。

[0099] ステップ S 2 7 までの処理でニュース情報を割り付けられた知見抽出対象語ごとに、知見抽出対象別判定部 1 5 3 は、以下に記載するステップ S 2 8 からステップ S 3 4 の処理を行う。

[0100] ステップ S 2 8 では、ステップ S 2 4 の第 2 概念、すなわち株価変動材料（表現）抽出結果を参照し、現在処理中の知見抽出対象語である銘柄が属する第 1 概念に特有の第 2 概念が抽出されているか否かを判定する。

[0101] 処理中の知見抽出対象語において、当該知見抽出対象語の属する第 1 概念に関連づいて第 2 概念が抽出されている場合、当該知見抽出対象語に関する情報が変動する可能性があると判断し、当該知見抽出対象語と第 2 概念名のセットをアラームに追加する（ステップ S 2 9）。

[0102] 具体的に記載すると、処理中の銘柄において、当該銘柄の所属グループに対応付けられた株価変動材料が抽出されている場合は、その銘柄の株価が変動する可能性があると判断し、その銘柄と株価変動材料のセットを株価変動材料アラームに追加する。

[0103] 記憶部 1 1 に保持された知見情報が「第 1 概念名・知見抽出対象語」と第 2 概念名に分けられている場合は、ステップ S 2 8 は、現在処理中の知見抽出対象語が属するか否かに関わらず記憶部 1 1 に保持された第 2 概念全てに

関して抽出されているか否かを判定する。

- [0104] この場合、ステップS 2 9では、知見抽出対象語である銘柄と、ステップS 2 8で抽出された第2概念である株価変動材料のセットを第2概念アラームに追加する。
- [0105] 知見抽出対象語である銘柄にニュース情報が割り付けられた後で、処理中の銘柄に割り付けられ、記憶部11に記憶された過去一定期間、例えば3日間のニュース見出し・ニュース本文を参照して急上昇ワードがあるかを判定し（ステップS 3 0）、該当銘柄のニュース情報に急上昇ワードが存在する場合は、銘柄と急上昇ワードのセットを急上昇アラームに追加する（ステップS 3 1）。
- [0106] 急上昇ワードは、近年、検索サイトなどで公開されているもので、出現単語の新しさ、急上昇を判定する技術が公知となっている。
- [0107] さらに、同様に処理中の銘柄に割り付けられ、記憶部11に記憶された過去一定期間、例えば1時間のニュース情報のニュース見出し・ニュース本文を参照して処理中の受信ニュース情報が過去一定期間のニュース情報と類似・一致しているかを判定し（ステップS 3 2）、類似ニュース情報もしくは一致するニュース情報がある場合は、そのニュース情報対の類似度を求め、多重配信アラームを追加する（ステップS 3 3）。
- [0108] 以上の処理により出力された第2概念アラーム、急上昇アラーム、多重配信アラームは、情報分析部15から出力され、アルゴリズムトレードエンジンなどのシステムもしくは株取引を行うユーザ端末装置2に提供される（ステップS 3 4）。
- [0109] 提供先がアルゴリズムトレードエンジンなどのシステムである場合、アルゴリズムトレードエンジンなどのシステムは、知見抽出装置1から提供されたニュース情報とニュース情報に付与された情報の他に、例えば該当銘柄の現在の株価、注文状況などを入力情報として、現時点で取るべき取引戦略を、株取引を行うユーザ端末装置2を介してユーザに提示する。
- [0110] 上述したように本実施形態によれば、受信されたニュース情報から経済用

語を含むニュース情報を選択し、当該選択されたニュース情報から知見情報を抽出し、当該抽出された知見情報を記憶する構成により、銘柄に関連付けられていない状態で随時配信されるニュース情報から銘柄グルーピングや株価変動材料の知見情報を随時抽出することにより、知見情報を最新の状態にすることができる。

[0111] また、株価変動評価語を含むニュース情報が銘柄の第1概念名を定義するグループ定義語を含むとき、当該グループ定義語に基づいて当該ニュース情報から第1概念名を抽出し、株価変動評価語を含むニュース情報から企業名及び証券コードを含む銘柄を抽出し、株価変動評価語を含むニュース情報から株価変動材料を抽出し、当該抽出された第1概念名、銘柄及び株価変動材料を組み合わせることで知見情報を作成し、当該作成された知見情報を記憶部11に書込む構成により、証券市場で使われる固定の業種ではないグルーピングの知見情報を活用できるので、株価変動予測を有効に支援できる。

[0112] さらに、知見情報が記憶された後、新規に受信された全てのニュース情報を記憶部11内の知見情報に基づいて分析する構成により、分析結果を、即時にアルゴリズムトレードエンジンなどのシステムもしくは株取引を行うユーザに提供することで、直近の株取引を有効に支援することができる。

[0113] また、提供先がアルゴリズムトレードエンジンなどのシステムである場合、現在の株価や注文状況などの数値情報のみを参照する場合と比較して、豊富な情報をもとにした判定を行うことが可能となる。

[0114] また、本実施形態及び以下の各実施形態は、対象をインターネット配信ニュースに限定しない。例えば、インターネットで配信される他の情報（ブログ、ミニブログ、マイクロブログ、企業発表情報）や、インターネット外に存在する電子文書を対象とすることも可能である。さらに、対象をニュース情報やブログなどの複数種類の情報の組み合わせとすることも可能である。

[0115] なお、ステップS32の類似ニュース情報もしくは一致するニュース情報の判定には、例えば、この出願の出願時に未公開である先願（特願2010-247518）の明細書に記載した処理[1]～[3]を用いてもよい。

以下の処理〔１〕～〔３〕において「タイトル」の語は「ニュース本文」と読み替えてもよい。また、「ＩＤ付ニュース情報」の語は「ニュース情報」と読み替えてもよい。また、処理〔１〕～〔３〕の主体は、前述したステップＳ３２に合わせて「知見抽出対象別判定部１５３」に書き換えている。

[0116] 〔１〕記憶部１１は、単語解析辞書及び同一性判定基準などを予め記憶する。

[0117] 同一性判定基準は、送信されるニュース情報における任意の２件のニュース情報が同一ニュースであるか否かを判定する基準を示しており、当該２件のニュース情報に含まれるニュースソース名が互いに一致することと、当該２件のニュース情報に含まれる配信日時の差分を示す配信時間差が基準値（最大の配信時間差）よりも小さいことと、当該２件のニュース情報における２件のタイトルの形態素解析結果から算出される類似度が規定値より高いことと、当該類似度は前記数値情報の有効数字の桁を四捨五入により合わせた後に算出されることとを含んでいる。

[0118] なお、類似度は、例えば、２件のタイトルの形態素解析結果に含まれる自立語の集合全体における当該自立語及び数値情報の個数に対し、当該２件のタイトルの形態素解析結果の両方に含まれる自立語及び数値情報の個数が占める割合である。また、類似度（０以上１以下）の規定値は、０．９程度の高い値が好ましい。また、「２件のタイトルの形態素解析結果から算出される類似度が規定値より高いこと」に代えて、「２件のタイトルの形態素解析結果から抽出される自立語が互いに一致すること」としてもよい。また、同一性判定基準は、同一性判定ルールと読み替えてもよい。

[0119] 〔２〕知見抽出対象別判定部１５３は、記憶部１１内の単語解析辞書を用い、ＩＤ付ニュース情報の単語解析処理を行なう。単語解析処理は例えば一例として、形態素解析技術（公知の技術）を用いる。換言すると、知見抽出対象別判定部１５３は、記憶部１１内のＩＤ付ニュース情報に含まれるタイトルを形態素解析し、得られた形態素解析結果を当該ＩＤ付ニュース情報のニュースＩＤ及びタイトルに付加し、得られた解析結果情報を記憶部１１に

書き込む形態素解析機能をもっている。

[0120] なお、以下では形態素解析技術を一例として説明するが、知見抽出対象別判定部 153 での処理は、単語解析辞書を用いない、N グラムなどの形態素解析ではない解析処理を用いて、単語に分解してもよい。つまり、本実施形態は、形態素解析に限定しない手法によりニュース情報を単語に分割し、単語の比較で類似度を判断するものである。

[0121] 但し、例えば文字を単位とした N グラムの場合、品詞付けや自立語であるか否かの判定はできない。従って、形態素解析の他の手段で単語解析処理を行なうときは、「自立語」ではなく「単語」が処理対象となり得る。

[0122] [3] 知見抽出対象別判定部 153 は、記憶部 11 内の同一性判定基準を満たすか否かに基づいて、記憶部 11 内の最新の ID 付ニュース情報と過去に書き込んだ ID 付ニュース情報とが同一ニュースであるか否かを判定する。

[0123] 例えば、知見抽出対象別判定部 153 は、最新の ID 付ニュース情報におけるタイトルの形態素解析結果から抽出された自立語及び数値情報と、過去に書き込まれた ID 付ニュース情報におけるタイトルの形態素解析結果から抽出された自立語及び数値情報とが一致する割合を示す類似度を算出する。但し、類似度を算出する前に、数値情報の有効数字の桁を四捨五入により合わせておく。この類似度が規定値（例、0.9）より高い場合、同一性判定部 15 は、ID 付ニュース情報内のニュースソース名が一致し、さらに、各ニュースの配信時刻の差が基準値（例、5 分）以内ならば、同一ニュースである旨を判定する。なお、類似度が高い旨の確認、ニュースソース名の一致確認、配信時刻の差の確認は、任意の順序で実行可能である。また、類似度が高い旨の確認に代えて、自立語が完全一致する旨を確認してもよい。

[0124] 続いて、同一性判定の処理 [3] について、ニュース ID “38” のニュース情報と同一性判定対象のニュース ID “3” のニュース情報とを例に挙げて述べる。

[0125] ニュース ID “38” を含む最新の ID 付ニュース情報に対する形態素解

析結果からは、数値情報「15.8%減」1種、自立語「米<名詞-固有名詞-国>」「商品販売<名詞-一般>」「A社<名詞-固有名詞-組織>」「1月<名詞-副詞可能>」「リコール<名詞-サ変接続>」「問題<名詞-ナイ形容詞語幹>」「響く<動詞-自立>」7種が抽出される。

[0126] 同一性判定対象のニュースID“3”を含む過去のID付ニュース情報に対する形態素解析結果からは、数値情報「15%減<数値情報>」1種、自立語「A社<名詞-固有名詞-組織>」「1月<名詞-副詞可能>」「米<名詞-固有名詞-国>」「商品販売<名詞-一般>」4種が抽出される。

[0127] ここで、ニュースID“38”とニュースID“3”に対応する数値情報はそれぞれ「15.8%減」と「15%減」であり、値が一致しない。

[0128] 自立語については、2件の形態素解析結果から抽出される自立語が全部で7種あるのに対して、2件の形態素解析結果の両方から抽出される自立語が4種である。

[0129] この場合、数値情報1種と自立語全7種の合計8種の情報のうち一致するものが4種であることに基づき、一致度が50%、類似度が0.5と算出される。

[0130] 類似度の算出は、さらに「数値情報が含まれ、一致しない場合は類似度0とする」又は「数値情報の一致度と、自立語の一致度の平均を類似度とする」などとして実行してもよい。

[0131] なお、知見抽出対象別判定部153による自立語比較は、形態素解析結果をそのまま比較したが、これに限らず、形態素解析結果における動詞・形容詞・形容動詞を原形に変換して比較する処理や、否定の助動詞が続いていた場合には原形に戻す際に否定形の終止形とする処理、などの処理を加えるように変形してもよい。

[0132] (第2の実施形態)

次に、第2の実施形態について説明する。なお、第2の実施形態は、株取引に関する知見情報を抽出する第1の実施形態の知見抽出装置1に関する変形例である。

- [0133] 株価変動の方向は、新たに得られた情報がポジティブかネガティブであるかは必ずしも一致しない。株取引を行うユーザたちが事前に予測していた内容と比較して改善方向か改悪方向かが株価変動の方向を決める傾向がある。
- [0134] したがって、第1の実施形態では、情報分析部15の処理において株価変動材料の有無とその表現のみをCEP (Complex Event Processing: 複合イベント処理) もしくはユーザへの提供内容とした。しかし、一部の株価変動材料は、株価変動に一定の方向性を与える傾向がある。
- [0135] そこで、第2の実施形態では、第2手掛かり語のうち、方向性が明らかであるものには、知見抽出部14内の材料抽出部145が、方向(ポジティブもしくはネガティブ)と、その強度を付与する形態としている。この強度は、例えば整数-5~5の範囲で予め設定される。
- [0136] これに伴い、第2概念抽出部145は、図4に示した第2手掛かり語辞書113に代えて、図16に示す第2手掛かり語辞書113aを用いる形態としている。
- [0137] 具体的には、第2概念抽出部145は、知見抽出判定語チェック部141による判定の結果、電子文書であるニュース情報が株価変動評価語を含むとき、当該ニュース情報から株価変動材料を抽出する前述した機能に加え、当該抽出された株価変動材料に基づいて、記憶部11内の第2手掛かり語辞書113aを検索し、当該検索された株価変動の方向及び強度を表す各情報を当該株価変動材料に付与する機能を更に備えている。
- [0138] なお、第2手掛かり語辞書113aは、手掛かり語を識別するコードと、株価変動材料の表現に対する手掛かり語と、株価変動の方向を表す情報と、株価変動の強度を表す情報とが関連付けられて記述されている。
- [0139] 次に、以上のように構成された知見抽出装置1の動作を説明する。なお、第1の実施形態の知見抽出装置1と同様の構成・動作については説明を省略する

ステップS 1～S 6までの動作は、前述同様に実行される。

[0140] 第2概念抽出部（以下、材料抽出部と記載する）145は、知見抽出判定語チェック部（以下、株価変動評価語チェック部と記載する）141による判定の結果、ニュース情報が株価変動評価語を含むとき、当該ニュース情報から株価変動材料を抽出する。具体的には、第2概念抽出部（以下、材料抽出部と記載する）145は、株価変動評価語を含むニュース情報のニュース見出しもしくはニュース本文に、第2手掛かり語辞書（以下、材料表現手掛かり語辞書と記載する）113a内の手掛かり語の1つ以上が含まれるか否かを判定する（ステップS 7）。

[0141] ステップS 7およびステップS 8の処理は、図9のフローチャートではステップS 6の後としているが、ステップS 3の処理の前、もしくはステップS 3からステップS 6の処理と並列で実施してもよい。

[0142] 材料抽出部145は、受信したニュース情報のニュース見出しもしくはニュース本文に、材料表現手掛かり語辞書113a内の手掛かり語が含まれる場合、その手掛かり語に係り受けしている名詞句を株価変動材料として抽出する（ステップS 8）。

[0143] 例えば図17に示すニュースID“4”からの知見抽出処理において、材料表現手掛かり語辞書113aに登録された手掛かり語「嫌気」を抽出し、手掛かり語「嫌気」に係り受けしている名詞句「業績下方修正」からなる株価変動材料に、手掛かり語「嫌気」に対応づけられた方向及び強度を付与し、株価変動材料「材料：業績下方修正 方向：ネガティブ 強度：2」が得られる。

[0144] 組み合わせ作成部146は、前述同様に、知見情報を作成し、当該知見情報を記憶部11に書き込む（ステップS 9）。図18は当該知見情報の一例である。ここで、知見情報は、方向及び強度が付与された株価変動材料「材料：業績下方修正 方向：ネガティブ 強度：2」を含んでいる。知見情報は、第1の実施形態と同様に、組み合わせを、第1概念名、銘柄、株価変動材料をセットにしているが、組み合わせは第1概念名と銘柄のみとし、株価変

動材料は単独で知見情報として保持するとしてもよい。または、株価変動材料は第1概念名ではなく銘柄とセットにして知見情報として保持するとしてもよい。

[0145] その後、図17に示すニュースID“5”を受信した際は、情報分析部15は、ニュース本文から株価変動材料「業績下方修正」を抽出した際に、「方向：ネガティブ 強度：2」の情報を、銘柄「1152：AA製鉄（株）」材料「業績下方修正」に加えてアルゴリズムトレードエンジンなどのシステムもしくは株取引を行うユーザ端末装置2に提供する。

[0146] 上述したように本実施形態によれば、株取引に関する知見情報を抽出する知見抽出装置1によって抽出された株価変動材料に基づいて、第2手掛かり語辞書を検索し、当該検索された株価変動の方向及び強度を表す各情報を当該株価変動材料に付与する構成により、第1の実施形態の効果に加え、株価変動材料に方向と強度を加えた情報を提供することで、アルゴリズムトレードエンジンなどのシステムもしくは株取引を行うユーザの速やかな判断をさらに強く支援することができる。

[0147] （第3の実施形態）

次に、第3の実施形態について説明する。

[0148] 第3の実施形態は、株取引に関する知見情を抽出する第1の実施形態の知見抽出装置1の変形例であり、知見抽出部14が抽出する知見情報のある対象に関する材料（イベント）とそのタイムスタンプとしている。

[0149] これに伴い、情報受信部12は、ニュース情報を受信する前述した機能に加え、ニュース情報を受信する際に、当該ニュース情報にタイムスタンプを付加する機能を更に備えている。

[0150] 知見抽出部14は、情報選択部13により選択されたニュース情報から知見情報を抽出する前述した機能において、当該選択されたニュース情報から所定の対象を示す表現を含むイベント情報を抽出すると共に、当該イベント情報が抽出されたニュース情報から抽出し、当該抽出したイベント情報及びタイムスタンプを知見情報として記憶部11に書込む機能をもっている。

- [0151] ここで、所定の対象を示す表現は、例えば、銘柄の企業名を表す企業名表現、又は銘柄の第1概念名を定義するグループ定義語の直前の名詞を表す表現である。なお、「銘柄の企業名を表す企業名表現」は、知見抽出対象語辞書（以下、銘柄辞書と記載する）112に記述されている。「銘柄の第1概念名を定義するグループ定義語」は、第1手掛かり語辞書（以下、グループ定義後辞書と記載する）111に記述されている。
- [0152] 次に、以上のように構成された知見抽出装置1の動作を説明する。図19は、知見抽出部14の処理動作の一例を示すフローチャートである。
- [0153] 情報受信部12がニュース情報を受信すると（ステップS41）、知見抽出部14はグループ定義語辞書111及び銘柄辞書112などを参照し、ユーザによって選択されている対象（企業が関連づけられた第1概念）を示す表現が含まれるか否かを判定する（ステップS42）。
- [0154] ユーザによって選択されている対象を示す表現が含まれる場合、知見抽出部14は、その対象を示す表現を含む名詞句を、第2概念（イベント）として抽出する（ステップS43）。
- [0155] 知見抽出部14は、抽出した第2概念を、対象名、受信情報のタイムスタンプと合わせて記憶部11に追加的に書込む（ステップS44）。
- [0156] 以下、情報分析部15は、前述同様に、ステップS21～S34の動作を実行する。
- [0157] 図20は、知見情報の一例を示す図である。図5に示すニュースID“1”，“2”から第1概念名「インフルエンザ」に関して処理を行った例である。
- [0158] 抽出する知識（第2概念）はユーザによって選択された対象表現を含む名詞句としたが、この名詞句中から複合名詞を選択したり、固有名詞（地名など）を選択したり、急上昇ワードを選択して第2概念とし、第2手掛かり語を「〇県」「集団感染」「△市」やその組み合わせとしてもよい。
- [0159] 第1概念「インフルエンザ」で抽出した第2概念を、図12に示した如き、第1の実施形態で取得した知見情報にしたがって企業に展開し、図20に

示す知見情報で「対象」を「インフルエンザ」から「１０３３：y y ボウホールディングス（株）」に置き換えた情報を合わせて知見情報として書き込んでもよい。

[0160] さらに、図２０に示すように、類似度の高いニュース情報の配信数を「類似情報数」として合わせて知見情報としてもよい。

[0161] 上述したように本実施形態によれば、ニュース情報を受信する際に、ニュース情報にタイムスタンプを付加し、選択されたニュース情報から所定の対象を示す表現を含むイベント情報を抽出すると共に、当該抽出したイベント情報及びタイムスタンプを知見情報として記憶部１１に書込む構成により、第１の実施形態の効果に加え、ある対象（第１概念）に関する第２概念をタイムスタンプ情報のある時系列で抽出し保持することで、同様の事象が発生した際に、近い将来の予想に有効な知見情報を保持しユーザ端末装置２に提示することができる。

[0162] なお、第３の実施形態は、受信する情報を電子メールとし、対象をプロジェクトとすることで、企業活動のうちのプロジェクトに関するリスクを示す知見情報を抽出することも可能である。さらに、受信する情報を情報機器のシステムログとし、対象をコンピュータネットワークシステムとすることで、コンピュータネットワークシステムの障害検知に関する知見情報の抽出をすることも可能である。

[0163] （第４の実施形態）

次に、第４の実施形態について説明する。

[0164] 図２１を参照して、第４の実施形態の知見抽出装置１００の構成について説明する。なお、第１の実施形態乃至第３の実施形態に記載の構成については省略する。

[0165] 図２１に示すように、本実施形態の知見抽出装置１００は記憶部１１の評価情報１１４と確信度更新ルール１１５、及び知見評価更新装置１８を備える。評価情報１１４と確信度更新ルール１１５、及び確信度については後述する。

- [0166] 知見評価更新装置 18 は、上位概念チェック部 181 と、下位概念抽出部 181 と、評価部 183 と、確信度更新部 184 と、知見更新部 185 を備える。
- [0167] 上位概念チェック部 181 は、知見情報に含まれるグループ名毎に、あらかじめ設定した期間に記憶部 11 に新規登録されたニュース情報中で、当該グループ名が記載されたニュースの有無と配信数をチェックする。すなわち、本実施形態の知見情報は階層構造の情報であり、グループ名を上位概念とし、銘柄を下位概念とする。
- [0168] 下位概念抽出部 182 は、上位概念チェック部 182 によってチェックされたグループ名の下位概念である、知見情報に含まれる銘柄を全て抽出する。また、下位概念抽出部 182 は、上位概念チェック部 181 がチェックしたニュース情報から当該銘柄が記載されたニュースの有無と配信数をチェックする。
- [0169] 評価部 183 は、記憶部 11 に記憶された評価情報 114 に基づいて、下位概念抽出部 182 が抽出した銘柄毎に評価値を算出する。評価情報 114 は、例えば株価情報、ニュース、株の出来高、会社の業績や従業員数などである。
- [0170] 確信度更新部 184 は、評価部 183 が算出した評価値と、確信度更新ルール 15 とに基づいて、後述する確信度を更新する。
- [0171] 知見更新部 185 は、確信度更新部 184 による更新結果に基づいて、記憶部 11 に記憶された知見情報を更新する。
- [0172] ここで、株価変動が、各銘柄が関連するニュースの発生の影響を受けることは前述のとおりであるが、影響の大きさは銘柄や、グループ名（キーワード）により異なる。したがって、グループ名と銘柄の組み合わせに加えて、その銘柄が該当グループ名のニュース発生時に受ける影響の強さ、すなわち、その銘柄が該当グループに所属する確信度の情報を合わせて保持することが有効である。
- [0173] したがって、第 4 の実施形態では、図 12 に示した知見情報に代えて、図

22に示すように、確信度と確信度更新履歴とを含む確信度情報を加えた知見情報を保持する形態としている。すなわち、本発明の確信度とは、下位概念が上位概念に帰属する（関連する）度合いを示す数値であり、背景技術に記載した第3の技術における確信度とは異なる。知見情報における確信度が高いほど、下位概念が上位概念に帰属する度合いが高い、すなわち関連性が強い情報であると判定できる。

[0174] なお、図22では図12の株価変動材料の情報を省略しているが、株価変動材料を加えた構成としてもよい。または、グループ名と株価変動材料の組合せを図22の知見情報とは別テーブルで管理する構成としてもよい。さらに、株価変動材料に加えて、第2の実施形態および図18に示すように、方向、強度を合わせて管理してもよい。また、確信度情報に含まれる情報は確信度のみでもよい。

[0175] 図22の知見情報では、第1の実施形態、第2の実施形態、第3の実施形態で用いた図12の知見情報と同様にグループ名、銘柄の2種の情報を保持する。さらに、その銘柄が該当グループに所属する確信度の情報を加えた3種の情報を必須とする。これらに加えて、グループ名を細分化する下位グループ名、該当企業から発表された、該当事業（グループ）に参入予定の時期である事業参入時期、該当企業から発表された、該当事業から撤退予定の時期である事業撤退時期、過去一定回数までの確信度更新履歴、そのグループ名と銘柄の関係を新規抽出した際のグループ定義語を記録するグループ定義語、などの情報を合わせて保持してもよい。本実施形態では、確信度更新履歴を2つ前の履歴まで保持する。

[0176] 次に、図22の知見情報の確信度を保持し更新する知見評価更新装置18を備えた知見抽出装置100の動作を説明する。図23は、知見評価更新装置18の処理動作（以下、知見評価更新処理という）の詳細を示す模式図、図24及び図25は知見評価更新処理の一例を説明するためのフローチャートである。

[0177] 以下、本実施形態の知見抽出装置100において、図22の4行目の知見

情報に関する処理を行う場合について説明する。図 2 2 の 4 行目の知見情報は、「第 1 概念名：インフルエンザ、下位グループ名：インフルエンザ薬、銘柄：f f 医薬<1 0 4 2>、確信度：1、事業参入時期：2 0 0 7 年 1 0 月 1 9 日、確信度更新履歴 1（前回の確信度更新履歴）：2 0 0 8 年 9 月 1 0 日 + 1、グループ定義後：関連」、である。

[0178] 第 1 概念名（以下、グループ名と記載する）、銘柄（企業名表現）、株価変動材料のセットの抽出は、第 1 の実施例と同様に実行される。グループ名と銘柄の組み合わせが新規に抽出された際は、確信度の初期値として、例えば 1. 0 を付与する。また、本実施形態の確信度の最小値は 0 とし、後述する確信度更新処理により、確信度が 0 を下回る結果となった場合は 0 に修正する。また、本実施形態の確信度の最大値は 5. 0 とする。確信度更新処理により、確信度が 5. 0 を超える結果となった場合は、例外処理により、5. 0 を大きく超えることがないように調整してもよい。例外処理は、例えば、前日の確信度が 5. 0 より大、かつ、今回の更新でも確信度が増える状況の場合は、確信度を 0. 0 5 のみ増やすこととし、前日の確信度が 4. 9 5 未満、かつ今回の更新で確信度が 5. 0 を超える場合に、更新後の確信度を 5. 0 とする、という処理を行う。

[0179] なお、以下の説明では、知見評価更新処理を、証券市場終了後の夕方もしくは夜間に 1 回実行するとしているが、この処理は、1 日 1 回ではなく、証券市場の前場と後場の終了時にそれぞれ実施してもよい。

[0180] 知見評価更新処理が起動されると、上位概念チェック部 1 8 1 は図 2 2 の知見情報からグループ名を 1 つ選択する（ステップ S 5 1）。ここでは、グループ名「インフルエンザ」が選択される。続いて、上位概念チェック部 1 8 1 は、その日に記憶部 1 1 に新規登録されたニュース中で、ステップ S 5 1 で選択されたグループ名が記載されたニュースの有無を確認する（ステップ S 5 2）。

[0181] なお、本実施形態では、ニュース件数の集計は、例えば営業日の 1 5 : 0 0 などの所定の時刻ごとに行う。休日に配信されたニュースは、翌営業日の

ニュースに加算してもよい。また、ステップS 5 2においては、グループ名と下位グループ名の両方が記載されたニュースの有無を確認してもよい。

[0182] 選択中のグループ名が記載されたニュースがない場合（ステップS 5 2がN o）、ステップS 5 3～ステップS 6 3の処理を省略し、ステップS 6 4に進む。ステップS 6 4については後述する。

[0183] 選択中のグループ名が記載されたニュースがある場合（ステップS 5 2がY e s）、下位概念抽出部1 8 2は、図2 2の知見情報のうち、グループ名が選択中のものと同一である知見情報に含まれる銘柄を抽出する（ステップS 5 3）。評価部1 8 3は、下位概念抽出部1 8 2によって選択された銘柄から1つを選ぶ（ステップS 5 4）。ここでは銘柄「f f 医薬<1 0 4 2>」が選択される。

[0184] 続いて、評価部1 8 3は、下位概念抽出部1 8 2によって選択された銘柄と、その銘柄に対応するグループ名と同時に含み、かつ事業撤退に関連するあらかじめ定めた語句（「撤退」「売却」など）がないニュースがあるか否かを判定する（ステップS 5 5）。

[0185] 下位概念抽出部1 8 2によって選択された銘柄と、その銘柄に対応するグループ名と同時に含み、かつ事業撤退に関連する語句がないニュースがない場合（ステップS 5 5はN o）、ステップS 5 6、S 5 7の処理を省略し、ステップS 5 8に進む。

[0186] 下位概念抽出部1 8 2によって選択された銘柄と、その銘柄に対応するグループ名と同時に含み、かつ事業撤退に関連する語句がないニュースがある場合（ステップS 5 5はY e s）、評価部1 8 3は、当該グループ名を含むニュース件数が規定数以上かどうかを確認する（ステップS 5 6）。本実施形態ではステップS 5 6での規程数を3とする。

[0187] 当該グループ名を含むニュース件数が規定数以上の場合（ステップS 5 6がY e s）、評価部1 8 3は、評価値として、当該銘柄の「出来高変化率」を算出し、算出した出来高変化率が後述する確信度の更新条件に合致するかどうかを確認する（ステップS 5 7）。

[0188] この場合、本実施形態の評価情報 114 には株の銘柄に関する日ごとの出来高が含まれる。銘柄が $f + \text{医薬} < 1042$ の出来高の一例を図 26 に示す。

[0189] ストップ 57 で評価部 158 は、この評価情報 114 を用いて「出来高変化率」を以下の式で算出する。

[0190] (1) $\text{出来高変化率} = (\text{当日の出来高}) / (\text{直近 5 営業日の出来高の平均})$

上記の式 (1) では、直近 5 営業日の平均を用いているが、5 営業日に限らないことも言うまでもない。

[0191] なお、ストップ 57 における更新条件は例えば「あらかじめ設定した閾値以上であれば更新する」である。また、更新条件は、当該知見抽出装置 100 に含まれる記憶部 11 に含まれる情報に基づいて動的に変化させてもよい。更新条件の動的な変化とは、例えば、評価情報 114 としてニューズ情報をを用いる場合、ニューズ情報に含まれる当該銘柄の関連ニュースの件数が所定の数よりも多い場合に、更新条件に含まれる閾値を所定の単位で上昇させるようなことが考えられる。

[0192] 本実施形態では、更新条件は「出来高変化率 > 1 、1 もしくは出来高変化率 < 1 、0 に当てはまるか否か」とする。

[0193] ストップ 55 ~ ストップ 57 の判定は、確信度を更新するかを判定するための処理の一例であり、判定をさらに詳細に分けることや、判定基準として別の指標を用いることも考えられる。

[0194] 当該グループ名を含むニューズ件数が規定数未満の場合 (ストップ 56 が N)、もしくは評価部 183 が算出した評価値が、確信度の更新条件に合致しない場合 (ストップ 57 が N)、ストップ 58 ~ ストップ 56 の処理を省略し、ストップ 63 の処理に進む。

[0195] 評価部 183 が算出した評価値が、確信度の更新条件に合致する場合 (ストップ 57 が Yes)、確信度更新部 184 が当該銘柄と対応付けて記憶部 11 に記憶された確信度を更新する (ストップ 58)。なお、確信度更

新部 184 による確信度の更新は、記憶部 11 に確信度更新ルール 115 をあらかじめ設定しておき、このルールに従って更新する。

[0196] 本実施形態における確信度更新ルール 115 の一例を図 27 に示す。図 27 に示す確信度更新ルール 115 は、記憶部 11 に格納されている。

[0197] 図 27 の確信度更新ルール 115 は、「ステップ S55 の判定が Yes、かつ出来高変化率 ≤ 1.1 の場合、確信度に 0.05 を加える」、「ステップ S55 の判定が Yes、かつ出来高変化率 > 1.1 の場合、確信度に (出来高変化率 $- 1.0$) $\times 0.5$ を確信度に加える」、「ステップ S56 が Yes、かつ出来高変化率 > 1.1 の場合、(ニュース変化率 $-$ 出来高変化率) $\times 0.5$ を確信度から減らす」、「ステップ S56 が Yes、かつ出来高変化率 < 1.0 の場合、(ニュース件数変化率 $-$ 出来高変化率) $\times 0.5$ を確信度に加える」である。

[0198] ニュース件数変化率は、以下の式に基づいて算出する。

[0199] (2) ニュース件数変化率 = (当日の該当グループ名を含むニュース件数に当日の確信度を乗じたもの) $\div$ (直近 5 営業日の該当グループ名を含むニュース件数に該当日の確信度を乗じたものの平均)

また、本実施形態におけるインフルエンザが含まれるニュースの件数を日ごとに示したグラフを図 28 に示す。

[0200] 本実施形態では、ステップ S55 の Yes からステップ S58 に進んだ場合と、ステップ S57 の Yes からステップ S58 に進んだ場合で更新ルールが異なる。また、例えば「出来高変化率 < 1.0 の場合 (1 $-$ 出来高変化率) $\times 0.5$ を前営業日の確信度から減らす」というような更新ルールでもよい。なお、確信度の前日値は例えば図 22 の知見情報に含まれる確信度更新履歴から算出する。または、確信度自体を履歴として保持してもよい。

[0201] 図 29 に、本実施形態の確信度に関するグラフの一例を示す。

[0202] 続いて、知見更新部 185 が、図 22 の知見情報を更新する。具体的には、知見更新部 185 は、更新後の確信度が、当該銘柄を含む知見情報を記憶部 11 から削除する条件 (以下、削除条件という) を満たすか否かを判定す

る（ステップS 5 9）。本実施形態では、例えば、確信度が0. 3以下になった場合は、知見更新部1 8 5は、削除条件を満たすと判定する。なお、この判定は、最新の確信度の数値で判定するのではなく、更新履歴で例えば減少が5回連続した場合に削除条件を満たすとするとしてもよい。

[0203] 更新後の確信度が0. 3より大きい場合（ステップS 5 9がN o）、すなわち、グループからの削除条件を満たさない場合、知見更新部1 8 5はステップS 6 1～ステップS 6 2の処理省略し、ステップS 6 3に進む。

[0204] 更新後の確信度が0. 3以下である場合（ステップS 5 9がY e s）、すなわち、グループからの削除条件を満たす場合、知見更新部1 8 5は、ユーザにアラームを提示する（ステップS 6 0）。このアラームとは、ユーザに選択中の銘柄を選択中のグループから削除するか否かを確認するためのものであり、例えば入出力インタフェース1 6に表示される。

[0205] ユーザが、表示されたアラームを参照し、当該知見情報を削除するか否かを、入出力インタフェースを用いて入力すると、知見更新部1 8 5は、入力された情報が削除を指示する情報であることを判定する（ステップS 6 1）。

[0206] 削除が指示されない場合（ステップS 6 1がN o）、知見更新部1 8 5はステップS 6 2の処理を省略し、ステップS 6 3の処理に進む。

[0207] 削除が指示された場合（ステップS 6 1がY e s）、知見更新部1 8 5は、選択された銘柄をグループから削除する（ステップS 6 2）。すなわち、選択中のグループと選択中の銘柄の組み合わせに該当する知見情報を記憶部から削除する。なお、銘柄のグループからの削除は、この一連の処理のみで行うのではなく、図2 2の知見情報に記載された事業撤退時期の情報に従い、撤退時期に到達した銘柄を削除する処理を別途行ってもよい。

[0208] 続いて、知見更新部1 8 5は、選択中のグループに未処理の銘柄があるか否かを判定する（ステップS 6 3）。未処理の銘柄がある場合（ステップS 6 3がY e s）、ステップS 5 4に戻って未処理の銘柄のうちの1つについて同様の処理を行う。

[0209] 選択中のグループで未処理の銘柄がなくなった場合（ステップS 6 3がN

o) もしくは、ステップS 5 2がN oの場合、知見更新部1 8 5は、知見情報に保持されたグループで未処理のものがあるか否かを判定する（ステップS 6 4）。

[0210] 未処理のグループがある場合（ステップS 6 4がY e s）、ステップS 5 1に戻って未処理のグループのうちの1 つについて同様の処理を行う。未処理のグループがない場合8ステップS 6 4がN o）、処理を終了する。

[0211] 上述したように本実施形態によれば、知見を抽出した情報源以外の情報（株価変動や出来高など）に基づいて、知見の有効性を評価し更新する構成により、第1の実施形態の効果に加え、知見の確かさを複数の情報源に基づき高く維持できる。

[0212] なお、第4の実施形態において、図2 4、図2 5の処理の判定基準は、前記に限定するものではない。例えば、ステップS 5 2の選択中のグループ名を含むニュースがあるか否かの判定は、選択中のグループ名を含むニュース数が前日以上であるか否かの判定に置き換え、もしくは選択中のグループ名が急上昇ワードに該当するか否かの判定に置き換えてもよい。

[0213] また、ステップS 5 7における選択中の銘柄の出来高変化率が更新条件を満たすか否かの判定は、選択中の銘柄の株価変動や売買代金、日経平均などの株式市場全体の傾向を示す指標と比較した株価変動傾向による判定に置き換えてもよい。

[0214] 例えば、ステップS 5 7において株価変動を用いて判定する場合、評価部1 8 3は、評価値としてその銘柄の直前の営業日の株価変動、例えばさらにその前の営業日からの株価変動の割合を求める。評価部1 8 3はこのとき、例えば株価変動が5 %以上の上昇ならば選択中の銘柄が選択中のグループに帰属することは適切であると判定する。また、評価部1 8 3は、株価変動が5 %未満の下落もしくは5 %未満の上昇の場合は、銘柄のグループ帰属の適切判断を保留する。また、評価部1 8 3は、株価変動が5 %以上の下落の場合は、選択中の銘柄が選択中のグループに帰属することは不適切であると判定する。この株価変動の判定は、第2の実施形態および図1 8に示したよう

に、ニュースに記載された材料表現とその方向も加味し、上昇（ポジティブ）と下落（ネガティブ）のいずれがグループに帰属することの判断に適切であるかを決定した上で行ってもよい。

[0215] また、各銘柄に関する処理は、所属するグループごとに実施する形態としたが、ある銘柄が同日に複数のグループで処理対象となる場合は、例外的な処理を実施してもよい。例外的な処理とは、例えば、[4] もしくは[5] の処理である。

[0216] [4] 同日に複数のグループで処理対象となる銘柄は、確信度更新を実施しない。

[0217] [5] 同日に複数のグループで処理対象となる銘柄は、所属するグループごとに、他の所属銘柄の変動を参照する。他の銘柄と同じ方向・同程度の数値変動をした場合は、そのグループでの確信度更新処理を実施する。

[0218] また、本実施形態では、確信度による評価更新の対象をグループ名と銘柄の関係としたが、確信度は、図22の知見情報に保持されたグループ定義語の評価に用いてもよい。すなわち、知見情報に保持されたグループ定義語と対応する確信度の情報をグループ定義語ごとに整理し、確信度の平均や最大値が規定値未満となったグループ定義語を有効ではないと判定して図2のグループ定義語辞書（第1手掛かり語辞書）から削除してもよい。

[0219] また、確信度の情報は、知見情報もしくはグループ定義語の評価更新に用いるものではなく、一般ニュースに含まれるグループ名を数える際に重み付けに用いてもよい。すなわち、例えばグループ名「インフルエンザ」を含むニュースが50件ある場合、確信度が1.0の銘柄には50件、確信度が0.8の銘柄には40件のニュース配信があったものとして件数を割り付けるようにしてもよい。

[0220] 図30は、本実施例の効果の一例を示す図である。図30には、ある東証一部上場銘柄について、その銘柄名を含むニュース件数（銘柄名のみ）、銘柄名もしくはその銘柄が所属するグループ名（図30の例では約200のグループがあるとする）の少なくとも一つを含むニュース件数（銘柄名ORグ

グループ名（確信度なし））、グループ名を含むニュース件数をそのグループ名の確信度で補正した件数と銘柄名を含むニュース件数の合計（銘柄名ORグループ名（確信度補正なし））について、ニュース件数と該当銘柄の出来高の相関係数を求めたものが示されている。

[0221] 相関係数とは、2つの確率変数の間の相関、すなわち類似度の度合いを示す統計学的指標であり、 -1 から $+1$ の間の実数値をとる。 1 に近いときは2つの確率変数には正の相関があるといい、 -1 に近ければ負の相関があるという。 0 に近いときはもとの確率変数の相関は弱い。相関係数と相関関係は、例えば相関係数が「 $0.0 \sim \pm 0.2$ 」であれば、相関関係は「ほとんど相関がない」、相関係数が「 $\pm 0.2 \sim \pm 0.4$ 」であれば、相関関係は「やや相関がある」、相関係数が「 $\pm 0.4 \sim \pm 0.7$ 」であれば、相関関係は「相関がある」、相関係数が「 $\pm 0.7 \sim \pm 0.9$ 」であれば、相関関係は「強い相関がある」、相関係数が「 $\pm 0.9 \sim \pm 1.0$ 」であれば、相関関係は「きわめて強い相関がある」、とみなせる。

[0222] 図30に示した、本実施形態の知見評価更新処理を適用した「銘柄名ORグループ名（確信度補正あり）」の相関係数は、銘柄名のみ、および銘柄名ORグループ名（確信度なし）の相関係数よりも高い値になっており、より強い相関がある知見情報を得ることができていると言える。

[0223] なお、本実施形態の知見評価更新装置18では、株取引に関する知見情報以外の知見情報を評価・更新することも可能であり、例えば商品に関する知見情報（以下、商品知見情報という）の知見評価更新処理が可能である。

[0224] この場合、商品知見情報は、企業の告知、ブログ、マイクロブログ、などのインターネットで配信される情報から抽出される。また、このとき、商品知見情報に含まれる下位概念の情報を商品名とし、上位概念の情報をグループ名とし、さらに商品知見情報の評価値を商品の売り上げ情報の変化率とする。この売り上げ情報の変化率により商品知見情報を評価・更新することで、商品のグループ名と売り上げとの相関がより強い知見情報を抽出することも可能である。

[0225] (第5の実施形態)

次に、第5の実施形態について説明する。

[0226] 第1の実施形態乃至第4の実施形態では、ニューズ情報にグループの情報があるかを判定する際に、グループ名のみを抽出すべき表現としていた。しかし、一般ニューズでは、経済ニューズで用いられるグループ名ではない関連語が記載される場合がある。例えば、図29に示すような経済ニューズからは、グループ名「防衛（関連）」、銘柄「a重工<7191>」、
「b重工業<7192>」、「c工業<7193>」の組み合わせが抽出される。なお、社名の横に示されている各数字は、架空の証券コードである。

[0227] 上記の場合、グループ名は「防衛」であるが、一般ニューズで報じられる内容は、図29の各経済ニューズの前半部分の「政府が自衛隊機の民間転用を進める方針を固めた」や「朝鮮半島情勢緊迫化」であり、グループ名「防衛」は記載されない可能性が高い。

[0228] そこで、第5の実施形態では、図12の知見情報に加えて、図32に示す関連語情報116を用いる。図32に示すように、本実施形態の関連語情報116は、知見N₀、とグループ名（第1概念名）と関連語とを対応付けたターズルとして記憶部11に保持される。なお、関連語はあらかじめユーザが登録してもよいし、ニューズ情報にグループ名と同時にでてくる単語を関連語として抽出して登録してもよい。

[0229] ここで、図33を参照して、第5の実施形態の知見抽出装置101の構成について説明する。第1の実施形態乃至第4の実施形態と同じ図番が付与されている各構成は、第1の実施形態乃至第4の実施形態と同じ動作を行なうものとする。

[0230] 図33に示すように、第5の実施形態の知見抽出装置101は、第4の実施形態の知見抽出装置100に加えて、記憶部11に関連語情報116を保持する。なお、本実施形態の知見抽出装置101は、第1乃至第3の実施形態の知見抽出装置116を加えた構成としてもよい。

[0231] 本実施形態の知見抽出装置 101 の処理の一例について、図 14、図 24 及び図 25 を参照して説明する。

[0232] 本実施形態の知見抽出装置 101 は、図 14 もしくは図 24 及び図 25 に示す処理において、ニュース情報に含まれるグループ名の有無を判定する際に（図 14 のステップ S24、もしくは図 24 のステップ S52）、図 32 に示す関連語情報 116 に含まれる関連語をグループ名と同等の表現としてチェック対象に加える。すなわち本実施形態の知見抽出装置 101 は、ステップ S51 で選択されたグループ名と同一のグループ名が関連語情報 116 に保持されているか検索し、保持されている場合、このグループ名もしくはこのグループ名に対応する関連語がニュース情報に含まれるか否かを判定する。

[0233] このとき、1つのニュースにグループ名と関連語が記載されている場合は、グループ名が記載されたニュース 1 件相当と数える。関連語は、第 1 の実施形態で説明した株価変動材料から、他のグループには登録されていない表現を選んでもよい。また、グループ名を含むニュース群と他のニュース群で出現頻度に有意な差のある名詞句を抽出してもよい。さらに、図 12 の知見情報と図 22 の知見情報と図 32 の関連語情報 116 は、グループ名で結合できる範囲で、別の分割形態で保持してもよい。

[0234] 本実施形態によれば、一般ニュースでは記載されることが少ないグループ名に加えて関連語を抽出対象とすることで、一般ニュースの銘柄への影響をさらに適切に評価することができる。

[0235] （第 6 の実施形態）

次に、第 6 の実施形態における知見抽出装置について説明する。第 6 の実施形態における知見抽出装置は、第 3 の実施形態と同様に、電子文書を受信する際に、当該電子文書にタイムスタンプを付加する機能を更に備え、知見抽出部 14 が抽出する知見情報のある対象に関する材料（イベント）とそのタイムスタンプとする。

[0236] 図 34 を参照して、第 6 の実施形態の知見抽出装置 102 の構成について説

明する。なお、上述の実施形態に記載の構成については省略する。

[0237] 図34に示すように、本実施形態の知見抽出装置102は知見評価更新装置19を備える。

[0238] 知見評価更新装置19は、図21に示した知見抽出装置100の知見評価更新装置18における、上位概念チェック部181に変えて電子文書検索部186を備え、下位概念抽出部182に変えて知見登録概念抽出部187を備える。

[0239] 電子文書検索部186は、知見情報に含まれる第1概念名に、あらかじめ設定した期間に記憶部11に新規登録された電子文書中で、当該第1概念名が記載された電子文書の有無と配信数をチェックする。

[0240] 知見登録概念抽出部187は、電子文書検索部186によってチェックされた第1概念名と関連付けられた、知見情報に含まれる知見抽出対象語を全て抽出する。また、知見登録概念抽出部187は、電子文書検索部186がチェックした電子文書から当該知見抽出対象語が記載された電子文書の有無と配信数をチェックする。

[0241] 第6の実施形態においては、電子文書をインターネットで配信されるマイクロブログ（ミニブログ、つぶやきブログとも称し、Twitter（登録商標）が代表例）の記事とし、第1概念をテレビ番組、知見抽出対象語をテレビ番組もしくはテレビ番組の登場人物もしくはテレビ番組の中のコーナーもしくはテレビ番組中で取り上げられる商品などの番組中に現れるもの、第2概念をテレビ番組もしくは知見抽出対象語に対する評価表現とする。すなわち、本実施形態では図21に示した第4の実施形態における知見抽出装置100のようにニュース情報から株取引に関する知見情報を抽出するのではなく、電子文書であるマイクロブログの記事のうち、テレビ番組に言及する内容のものを受信した際に、例えば、ドラマAに関連する知見抽出対象語（登場人物、取り上げられるもの、ストーリー展開、スタッフなど）をグルーピングする知識や、ドラマAもしくはその知見抽出対象語に関してよく使われる評価表現「ワロタ」「泣けた」「ばねえ」の知識を獲得し、これらの知

識を知見情報として抽出する知見抽出装置 102 について説明する。

[0242] なお、「ワロタ」とは「笑った」、もしくは「笑える」という意味で 사용되는語句であり、「ばねえ」とは、「半端なことではない」の意味で 사용되는語句である。また、知見抽出装置 102 は、マイクロブログを含むインターネット配信の電子文書、例えばニュース、一般ブログ、テレビ局発表情報を受信した際は、保持した知見情報を参照し、電子文書に「ドラマ A」などの番組名、該当番組に所属する知見抽出対象語、該当番組もしくは知見抽出対象語の評価表現の有無の情報を付与し、番組制作者などのユーザの端末装置に提供可能となっている。

[0243] すなわち、本実施形態の知見抽出装置 102 は、第 4 の実施形態に示した知見抽出装置 101 の変形例である。

[0244] 図 35 に、本実施形態における第 1 手掛かり語辞書 111 の一例を示す。図 35 に示すように、第 1 手掛かり語を識別するコードと、第 1 概念名、本実施形態においては番組名を定義する第 1 手掛かり語とが、関連付けられて記述されている。

[0245] 第 1 手掛かり語辞書 111 は、本実施形態においては、図 35 に一例を示すように、第 1 手掛かり語を識別するコードと、第 1 概念名、本実施形態においては番組名を定義する第 1 手掛かり語とが、関連付けられて記述されている。具体的には、番組名などの記事内容を定義する第 1 手掛かり語は、記事内容の先頭文字の「#」であるとする。

[0246] 本実施形態における知見抽出対象語辞書 112 は、図 36 に一例を示すように、知見抽出対象語、本実施形態においてはテレビ番組中に現れるものについて、知見抽出対象語を識別するコードと、知見抽出対象語が関連付けられた番組名と、知見抽出対象語のマイクロブログ記事中に現れる別表記である知見抽出対象語表現とが関連付けられて記述されている。知見抽出対象語表現としては、知見抽出対象語名の正称、略称、通称、俗称、愛称などの様々な名称が使用可能となっている。知見抽出対象語辞書 112 は、まず、テレビ局の公開情報に基づいて構築してもよい。その際、他の番組にも登場す

る出演者に関しては番組名を特定せず、番組に固有の役名などについては番組名を記載するとしてもよい。さらに、知見抽出対象語辞書は、知見情報を抽出する際に、第1概念が記載されたマイクロブログ記事に現れる別の名詞もしくは複合名詞を知見抽出対象語候補として記録し、ユーザの確認後に知見抽出対象語として追加登録するものとしてもよい。

[0247] 第2手掛かり語辞書113aは、図37に一例を示すように、第2手掛かり語を識別するコードと、第2概念、すなわち本実施形態においては知見抽出対象語表現に対す第2手掛かり語とが関連付けられて記述されている。本実施形態においては、第2手掛かり語は感情表現である顔文字およびそれに類する表現とする。第2手掛かり語辞書113aは、第2の実施形態と同様に、感情の方向と、その感情の強度が付与された形態とする。なお、強度は、感情が強いほど大きな値で示される。

[0248] 図38に、本実施形態で情報受信部12が受信する電子文書であるマイクロブログの記事の一例を示す。図38に示すように、本実施形態では、インターネットで配信されるマイクロブログ記事の構成要素のうち、図38に示す「タイムスタンプ」「ユーザ」「引用情報」「本文」の4項目と、情報受信部12が付与する「記事ID」を利用する。

[0249] 「タイムスタンプ」は、配信されたマイクロブログ記事に付与された発信日時情報である。「ユーザ」は、その情報を発信したユーザの識別名称である。「引用情報」は、他のユーザが発信したマイクロブログ記事をそのまま引用し再発信する際などに付与される、元の記事の発信ユーザの識別名称である。「記事本文」は配信されたマイクロブログ記事の本文の文字列である。

[0250] 情報選択部13は、情報受信部12が受信したマイクロブログの記事から、知見抽出対象とする電子文書、すなわち本実施形態ではテレビ番組に関するマイクロブログ記事を選択する。テレビ番組に関するマイクロブログであることの判断は、例えば、マイクロブログの記事内容の先頭に第1手掛かり語「#」に続いてテレビ番組名もしくはテレビ局名が記載されていることで

判断する。第 1 手掛かり語がないマイクロブログ記事の場合は、本文に、テレビ番組名、テレビ局名、もしくは図 3 6 の知見抽出対象語辞書 1 1 2 に記載された知見抽出対象語のいずれかが含まれることにより判断する。

[0251] 次に、以上のように構成された知見抽出装置 1 0 2 の動作について説明する。図 3 9 は知見抽出装置 1 0 2 における知見抽出部 1 4 の処理動作の詳細を示す模式図である。知見抽出部の処理動作は、第 1 の実施形態と同様に、図 9 のフローチャートに従う。

[0252] なお、以下の説明では、知見抽出部 1 4 の処理を、電子文書を受信する都度としているが、この処理は夜間などの情報分析処理を行わない時間帯に、受信済みのマイクロブログの記事に対してまとめて実施してもよい。

[0253] 情報受信部 1 2 が、マイクロブログの掲載されたサイトからの配信を受けるか、マイクロブログの掲載されたサイトにアクセスすることにより、新規のマイクロブログを受信する。例えば、図 3 8 に示す記事 ID “3” に該当するマイクロブログを受信する。

[0254] 情報選択部 1 3 は、情報受信部 1 2 が受信したマイクロブログの記事から、テレビ番組に関するマイクロブログ記事を選択する（ステップ S 1）。テレビ番組に関するマイクロブログであることの判断は、例えば、記事内容の先頭文字「#」に続いてテレビ番組名もしくはテレビ局名が記載されていることで判断する。記事内容の先頭文字「#」がないマイクロブログ記事の場合は、本文に、テレビ番組名、テレビ局名、もしくは図 3 6 の知見抽出対象語辞書に記載された知見抽出対象語のいずれかが含まれることにより判断する。選択されたマイクロブログ記事は、情報選択部 1 3 から知見抽出部 1 4 に送出される。

[0255] 知見抽出部 1 4 においては、知見抽出判定語チェック部 1 4 1 が、情報選択部 1 3 に選択されたマイクロブログ記事が知見抽出判定語を含むか否かを判定する（ステップ S 2）。本実施形態の場合は、ステップ S 1 の処理とステップ S 2 の処理の判断基準を同一とし、いずれかの処理ステップのみ実施するものとしてもよい。

- [0256] 知見抽出判定語が含まれる場合は、知見情報が含まれるマイクロブログ記事として、マイクロブログ記事を第１手掛かり語チェック部１４２に送信する。
- [0257] 第１手掛かり語チェック部１４２は、ステップＳ２の判定の結果、マイクロブログ記事が知見抽出判定語を含むとき、当該マイクロブログ記事が第１概念名、すなわち、本実施形態においては番組名を定義する第１手掛かり語を含むか否かを判定する（ステップＳ３）。
- [0258] 例えば、第１手掛かり語チェック部１４２は、第１手掛かり語辞書１１１を参照し、受信したマイクロブログ記事の本文に第１手掛かり語の少なくとも一つが含まれるか否かを判定する。例えば図３８に示す記事ＩＤ“３”のニュースは、第１手掛かり語「#」を含む。
- [0259] 受信したマイクロブログ記事に第１手掛かり語が含まれない場合は第１概念抽出部１４３によるステップＳ４の処理を省略する。
- [0260] 受信したマイクロブログ記事に第１手掛かり語が含まれる場合は、第１手掛かり語チェック部１４２は、マイクロブログ記事を第１概念抽出部１４３に送信する。
- [0261] 第１概念抽出部１４３は、ステップＳ３の判定の結果、マイクロブログ記事が第１手掛かり語を含むとき、当該第１手掛かり語に基づいて当該マイクロブログ記事から番組名を抽出する（ステップＳ４）。
- [0262] 例えば、第１概念抽出部１４３は、第１手掛かり語を含むマイクロブログの記事を受信すると、第１手掛かり語直後の名詞もしくは複合名詞を番組名として抽出する。受信したマイクロブログの記事が例えば図３８に示すニュースＩＤ“３”の場合は、第１手掛かり語「#」の直後の複合名詞「ドラマＡ」を番組名として抽出する。
- [0263] 本実施形態の場合は、例えば、番組名として手掛かり語直後の名詞もしくは複合名詞を抽出した後で、既存の番組名の一覧（図示せず）と比較し、抽出した名詞もしくは複合名詞が番組名の正式名称もしくは略称である場合のみ番組名として採用し、それ以外の場合は、番組名を含まないと判断しても

よい。

- [0264] 一方、知見抽出対象語抽出部 1 4 4 は、知見抽出判定語チェック部 1 4 1 による判定の結果、マイクロブログの記事が知見抽出判定語を含むとき、当該マイクロブログ記事から知見抽出対象語名を抽出する。
- [0265] 具体的には、知見抽出対象語抽出部 1 4 4 は、例えば、ステップ S 3 又は S 4 の後、第 1 概念抽出部 1 4 3 の処理を省略して送信されたマイクロブログ記事に対して、知見抽出対象語辞書 1 1 2 を参照し、マイクロブログ記事本文に知見抽出対象語表現が含まれるか否かを判定し（ステップ S 5）、知見抽出対象語表現が含まれる場合に当該知見抽出対象語表現に対応する知見抽出対象語を抽出する（ステップ S 6）。
- [0266] マイクロブログ記事本文に、知見抽出対象語辞書 1 1 2 の「知見抽出対象語表現」列の表現がある場合は、その行の「知見抽出対象語名」に記された知見抽出対象語が含まれると判断する。
- [0267] 知見抽出対象語辞書 1 1 2 は、一つの知見抽出対象語に対して複数行の定義もしくは知見抽出対象語表現列に複数の表現を含んでよい。また、1 件のマイクロブログ記事から複数の知見抽出対象語表現を抽出してよい。
- [0268] 例えば図 3 8 の記事 ID “3” のマイクロブログ記事の場合、知見抽出対象語表現「m さん」が含まれるため、知見抽出対象語「コード B O O 1 知見抽出対象語名 役名 M」が抽出される。
- [0269] 第 2 概念抽出部 1 4 5 は、知見抽出判定語チェック部 1 4 1 による判定の結果、マイクロブログ記事が知見抽出判定語を含むとき、当該マイクロブログ記事から第 2 概念、すなわち、本実施形態においては評価表現を抽出する。具体的には、第 2 概念抽出部 1 4 5 は、知見抽出判定語を含むマイクロブログ記事本文に、第 2 手掛かり語辞書 1 1 3 a 内の手掛かり語の一つ以上が含まれるか否かを判定する（ステップ S 7）。
- [0270] ステップ S 7 およびステップ S 8 の処理は、図 9 のフローチャートではステップ S 6 の後としているが、第 1 の実施形態と同様に、ステップ S 3 の処理の前、もしくはステップ S 3 からステップ S 6 の処理と並列で実施しても

よい。

[0271] 第2概念抽出部145は、受信したマイクロブログ記事の見出しもしくは本文に、第2手掛かり語辞書113a内の第2手掛かり語が含まれる場合、その手掛かり語と共起している語句を評価表現として抽出する（ステップS8）。

[0272] 例えば、図38に示す記事ID“3”の本文では、「mさん万能すぎワロタww#ドラマA」に第2手掛かり語「ww」が含まれる。

[0273] 第2手掛かり語「ww」と共起している語句、例えば直前に置かれた語句は「ワロタ」（一般辞書にはない未知語）である。この未知語「ワロタ」が評価表現として抽出される。

[0274] 評価表現は、処理中の1件の記事で手掛かり語直前の語句とせずに、同じ第2手掛かり語を含む多数のマイクロブログ記事から全ての単語を抽出し、該当の第2手掛かり語を含まない記事の集合と比較して統計的に共起する確率が高い語句を評価表現とするとしてもよい。例えば図38の記事ID7と記事ID8は、異なる第2手掛かり語「ww」および「（；__；）」の直前の語句が「ヤバイ」で一致し、どちらの第2手掛かり語と共起する確率も同じなので、評価表現としない。

[0275] 抽出された評価表現「ワロタ」には、第2手掛かり語辞書113aに登録された第2手掛かり語「ww」に対応づけられた方向及び強度を付与し、評価表現「評価表現：ワロタ 方向：笑い 強度：1」が得られる。

[0276] ステップS3からステップS8の処理の終了後、組み合わせ作成部146は、処理中のマイクロブログ記事から抽出された番組名、知見抽出対象語及び評価表現を組み合わせで知見情報を作成し、当該知見情報を記憶部に書き込む（ステップS9）。ここで、知見情報は、方向及び強度が付与された評価表現「評価表現：ワロタ 方向：笑い 強度：1」を含んでいる。評価表現は、第1の実施形態と同様に、方向や強度が付与されない形式を選択してもよい。

[0277] マイクロブログ記事から抽出された番組名が1つである場合は、その番組

名に、抽出された全ての知見抽出対象語、抽出された全ての評価表現を割り付ける。すなわち、1つの番組名に対して、知見抽出対象語と評価表現はそれぞれ1つまたは複数割り付けられることを許容する。抽出された番組名が複数である場合は、第1の実施形態と同様に、文脈解析や簡易的な判別処理により、番組名、番組名に係り受けしている知見抽出対象語、番組名もしくはその番組名に係り受けしている知見抽出対象語である銘柄に係り受けしている評価表現を判別して組み合わせてもよい。さらに、知見情報は番組名に対して知見抽出対象語と評価表現が割り付けられるとしているが、知見抽出対象語もしくは評価表現がない組み合わせが作成されてもよい。

[0278] このような処理により、番組名、知見抽出対象語名、評価表現のセットとして以下の知見が抽出され、記憶部11に書き込まれる。

[0279] 番組名 : ドラマA

 番組に属する知見抽出対象語 : 役名M

 評価表現 : ワロタ 方向 : 笑い 強度 : 1

組み合わせは、番組名、知見抽出対象語、評価表現をセットにするとしたが、第1の実施形態と同様に、組み合わせは番組名と知見抽出対象語のみとし、評価表現は単独で知見情報として保持するとしてもよい。または、評価表現は番組名ではなく知見抽出対象語とセットにして知見情報として保持するとしてもよい。

[0280] ステップS2で知見抽出判定語がないと判定されたマイクロブログ本文は、第1概念チェック部147により、マイクロブログ記事本文に、記憶部11に書き込まれた知見情報の番組名を含むか否かを判定される（ステップS11）。

[0281] 番組名を含む場合は、知見抽出対象語抽出部144が知見抽出対象語辞書112を用いて、マイクロブログ本文に知見抽出対象語表現が含まれるか否かを判定する（ステップS12）。

[0282] 知見抽出対象語表現が含まれる場合、例えば「採用」「登場」などの表現がある場合は、組み合わせ作成部146が抽出された知見抽出対象語表現に対

応する知見抽出対象語を番組に追加する処理のための組み合わせを作成する。

[0283] 「降板」などの表現がある場合は、組み合わせ作成部 146 が、番組に対して抽出された知見抽出対象語表現に対応する知見抽出対象語を取り除く処理のための組み合わせを作成する（ステップ S 13）。

[0284] ステップ S 13 で番組と知見抽出対象語の追加もしくは削除処理の組み合わせが作成された場合は、その処理を記憶部 11 に対して実行する（ステップ S 9）。

[0285] なお、知見抽出対象語の抽出は、知見抽出対象語辞書 112 を用いずに、番組名を含むマイクロブログ記事本文から名詞句を抽出し知見抽出対象語とするとしてもよい。この場合、抽出した名詞句は、新たに知見抽出対象語として知見抽出対象語辞書 112 に追加するものとしてもよい。さらに、ステップ S 11 から S 13 の処理と判断は、1 件のマイクロブログのみで行わずに、例えばある規定の数を超える発信があった場合にのみ記憶部 11 に対する実行対象としてもよい。または、ステップ S 11 から S 13 までの処理は、マイクロブログ本文に対して行うのではなく、テレビ局やニュース配信業者から配信された電子文書のみに対して実行するものとしてもよい。

[0286] 次に、情報分析部 15 の動作を図 39、図 14 及び図 15 によって説明する。

[0287] 始めに、情報受信部 12 は、マイクロブログの掲載されたサイトからの配信を受けるか、マイクロブログの掲載されたサイトにアクセスすることにより、新規のマイクロブログ記事を受信する。例えば、図 38 の記事 ID 6 に該当するマイクロブログ記事を受信する（ステップ S 21）。受信したマイクロブログ記事は知見抽出対象語・第 1 概念・第 2 概念抽出部 151 に送信される。

[0288] 知見抽出対象語・第 1 概念・第 2 概念抽出部 151 は、まず、記憶部 11 に記憶された知見抽出対象語辞書 112 を参照し、受信したマイクロブログ記事に知見抽出対象語表現があるか否かを判定する（ステップ S 22）。

[0289] 知見抽出対象語表現がある場合、知見抽出対象語・第 1 概念・第 2 概念抽

出部 1 5 1 は、抽出された知見抽出対象語表現に対応する知見抽出対象語名・コードを抽出する。また、知見抽出対象語・第 1 概念・第 2 概念抽出部 1 5 1 は、図 3 6 に示す知見抽出対象語辞書 1 1 2 を知見情報として参照し、抽出した知見抽出対象語名が割り付けられた第 1 概念、すなわち本実施形態においては番組名を全て抽出する（ステップ S 2 3）。知見抽出対象語・第 1 概念・第 2 概念抽出部 1 5 1 及び電子文書割付部 1 5 2 は、マイクロブログ記事本文から抽出された知見抽出対象語名ごとに、その知見抽出対象語が属する各番組に対して、ステップ S 2 4・ステップ S 2 5 の処理を行う。

[0290] 始めに、知見抽出対象語・第 1 概念・第 2 概念抽出部 1 5 1 は、これに先立つ処理によって抽出された、図 4 0 中に示す知見情報を参照し、その時点で処理中の知見抽出対象語名・その知見抽出対象語が属する番組名に割り付けられた第 2 概念、すなわち本実施形態においては評価表現を抽出し、マイクロブログ本文中に該当の評価表現があるか否かを判定する（ステップ S 2 4）。

[0291] 該当の評価表現がある場合は、処理中の受信マイクロブログ記事にその番組に関する情報と評価があると判断し、電子文書割付部 1 5 2 が、処理中のマイクロブログ記事を、知見抽出対象語及び番組に割り付ける（ステップ S 2 5）。本実施形態においては、第 1 の実施形態とは異なり、番組名に属する全ての知見抽出対象語への展開はしないとしてもよい。

[0292] 知見抽出対象語・第 1 概念・第 2 概念抽出部 1 5 1 は、ステップ S 2 2 からステップ S 2 5 の処理の後、もしくは、前、または並列で、受信したマイクロブログ記事本文に、図 1 2 に示す知見情報内の第 1 概念名が含まれるか否かを判定する（ステップ S 2 6）。

[0293] 例えば、図 3 8 に示す記事 ID “6” のマイクロブログ記事では、本文に番組名は含まれていない。この場合は、知見抽出対象語・第 1 概念・第 2 概念抽出部 1 5 1 は、抽出された知見抽出対象語「役名 M」が図 3 6 の知見抽出対象語辞書 1 2 で関連付けられた番組名「ドラマ A」を抽出して電子文書割付部 1 5 2 に送出する。

- [0294] 第1概念名が抽出された場合、電子文書割付部152は、図12に示す知見情報でその番組名に対応付けられた知見抽出対象語を取り出し、処理中のマイクロブログ記事を該当銘柄全てに割り付けることもできる（ステップS27）。本実施形態ではこの処理を省略し、マイクロブログ記事は、本文から抽出された知見抽出対象語のみに割り付けるものとする。
- [0295] ステップS27までの処理でマイクロブログ記事を割り付けられた知見抽出対象語ごとに、知見抽出対象別判定部153は、以下に記載するステップS28からステップS34の処理を行う。
- [0296] ステップS28では、ステップS24の第2概念、すなわち本実施形態では評価表現抽出結果を参照し、現在処理中の知見抽出対象語が属する番組の評価表現が抽出されているか否かを判定する。
- [0297] 処理中の知見抽出対象語で、処理中の所属番組に対応付けられた評価表現が抽出されている場合は、その知見抽出対象語もしくは番組への反応が変動する可能性があると判断し、その番組、知見抽出対象語と、評価表現のセットをアラームに追加する（ステップS29）。
- [0298] 記憶部11に保持された知見情報が「番組名・知見抽出対象語」と評価表現に分けられている場合は、ステップS28は、現在処理中の知見抽出対象語や番組が属するか否かに関わらず記憶部11に保持された評価表現全てに関して抽出されているか否かを判定する。
- [0299] この場合、ステップS29では、知見抽出対象語と、ステップS28で抽出された評価表現のセットをアラームに追加する。
- [0300] 知見抽出対象語にマイクロブログ記事が割り付けられた後で、処理中の知見抽出対象語もしくは番組に割り付けられ、記憶部11に記憶された過去一定期間、例えば3日間のマイクロブログ本文を参照して急上昇ワードがあるかを判定し（ステップS30）、処理中の知見抽出対象語もしくは番組のマイクロブログ記事に急上昇ワードが存在する場合は、番組、知見抽出対象語と急上昇ワードのセットを急上昇アラームに追加する（ステップS31）。
- [0301] さらに、同様に処理中の知見抽出対象語もしくは番組に割り付けられ、記

憶部 11 に記憶された過去一定期間、例えば 1 時間のマイクロブログ記事本文を参照して処理中の受信マイクロブログ記事が過去一定期間のマイクロブログ記事と類似・一致しているかを判定し（ステップ S 32）、類似マイクロブログ記事もしくは一致するマイクロブログ記事がある場合は、そのマイクロブログ記事対の類似度を求め、多重配信アラームを追加する（ステップ S 33）。

[0302] 以上の処理により出力されたアラーム、急上昇アラーム、多重配信アラームは、情報分析部 15 から出力され、ユーザ端末装置 2 に提供される（ステップ S 34）。

[0303] 上述したように本実施形態によれば、受信されたマイクロブログ記事からテレビ番組、知見抽出対象語（番組に取り上げられるものなど）、評価表現の知見を抽出して保持することで、一般の辞書にはない表現も知見として活用できる。

[0304] さらに、知見情報が記憶された後、新規に受信された全てのマイクロブログ記事を記憶部 11 内の知見情報に基づいて分析する構成により、分析結果を、即時にユーザに提供することで、視聴者の反応の把握を有効に支援することができる。

[0305] さらに、知見抽出部 14 は、情報選択部 13 により選択されたマイクロブログ記事から知見情報を抽出する前述した機能において、第 3 の実施形態と同様に、当該選択されたマイクロブログ記事から所定の対象を示す表現を含むイベント情報を抽出すると共に、当該イベント情報が抽出されたマイクロブログ記事から抽出し、当該抽出したイベント情報及びタイムスタンプを知見情報として記憶部 11 に書込む機能をもっている。

[0306] ここで、所定の対象を示す表現は、例えば、知見抽出対象語、すなわち本実施形態においては知見抽出対象語辞書 112 に記述された知見抽出対象語表現、又は第 1 概念、すなわち本実施形態においては番組名を表す表現である。なお、「知見抽出対象語の番組名を定義する第 1 手掛かり語」は、第 1 手掛かり語辞書 111 に記述されている。

- [0307] 知見抽出部 14 の処理は、第 3 の実施形態と同様に、図 18 のフローチャートに従う。
- [0308] 情報受信部 12 がマイクロブログ記事を受信すると（ステップ S 41）、知見抽出部 14 は第 1 手掛かり語辞書 111 及び知見抽出対象語辞書 112 などを参照し、ユーザによって選択されている対象（知見抽出対象語、番組名など）を示す表現が含まれるか否かを判定する（ステップ S 42）。
- [0309] ユーザによって選択されている対象を示す表現が含まれる場合、知見抽出部 14 は、その対象を示す表現を含む名詞句を、材料（イベント）として抽出する（ステップ S 43）。
- [0310] 知見抽出部 14 は、抽出した材料（イベント）を、対象名、受信情報のタイムスタンプと合わせて記憶部 11 に追加的に書込む（ステップ S 44）。
- [0311] 以下、情報分析部 15 は、前述同様に、ステップ S 21～S 34 の動作を実行する。
- [0312] 図 41 は、知見情報の一例を示す図である。図 38 に示すニュース ID “3”，“6” から知見抽出対象語名「役名 M」に関して処理を行った例である。
- [0313] 抽出する知識はユーザによって選択された対象表現を含む名詞句としたが、この名詞句中から複合名詞を選択したり、固有名詞を選択したり、急上昇ワードを選択して材料（イベント）としてもよい。
- [0314] 知見抽出対象語名「役名 M」で抽出した材料（イベント）を、図 36 の知見抽出対象語辞書 112 にしたがって番組に割り付け、図 41 に示す知見情報で「対象」を「役名 M」から「ドラマ A」に置き換えた情報を合わせて知見情報として書き込んでもよい。
- [0315] さらに、図 41 に示すように、類似度の高いマイクロブログ記事の配信数を「類似情報数」として合わせて知見情報としてもよい。
- [0316] 上述したように本実施形態によれば、マイクロブログ記事を受信する際に、マイクロブログ記事にタイムスタンプを付加し、選択されたマイクロブログ記事から所定の対象を示す表現を含むイベント情報を抽出すると共に、当

該イベント情報が抽出されたマイクロブログ記事から抽出し、当該抽出したイベント情報及びタイムスタンプを知見情報として記憶部 11 に書込む構成により、第 3 の実施形態と同様に、ある対象に関する材料（イベント）をタイムスタンプ情報のある時系列で抽出し保持することで、同様の事象が発生した際に、近い将来の予想に有効な知見情報を保持しユーザ端末装置 2 に提示することができる。

[0317] 続いて、知見評価更新装置 19 の処理について説明する。

[0318] 番組がコメディタッチである場合は「笑い」の評価は前向きととらえることができるが、番組がシリアスなドラマである場合は、「笑い」は失笑であり、低評価であることが推測される。したがって、番組ごとに、評価表現と、番組の評価の高さ、本実施形態では確信度の情報を合わせて保持することが有効である。

[0319] したがって、本実施形態では、図 42 に示すように、確信度と確信度更新履歴とを含む確信度情報を加えた知見情報を保持する形態としている。すなわち、本実施形態の確信度とは、第 2 概念と第 1 概念の関連する度合いを示す数値である。

[0320] なお、図 42 では第 1 概念である番組名、第 2 概念である評価表現、知見抽出対象語の情報を保持しているが、番組名もしくは知見抽出対象語のいずれかの情報は保持しないとしてもよい。

[0321] 知見評価更新装置 19 の処理は、本願の出願時点で未公開である特願 2012-15700 の処理に従って行うことができる。本実施形態の場合、数値情報として、例えば番組の視聴率を用いることができる。知見評価更新処理は、マイクロブログ記事の配信が少ない時間帯に 1 日 1 回の頻度で実施しても、分析対象としている番組終了から一定時間を経た後にそれぞれ実施してもよい。

[0322] 本実施形態によれば、知見を抽出した情報源以外の情報（視聴率など）に基づいて、第 1 概念と第 2 概念の組合せの有効性を評価し更新する構成により、知見の確かさや第 1 概念もしくは知見抽出対象語への影響度を複数の情

報源に基づき高く維持できる。

[0323] 確信度の情報は、知見情報もしくは第1手掛かり語もしくは第2手掛かり語の関連性の評価更新に用いるものではなく、マイクロブログ記事をはじめとする電子文書に含まれる番組名もしくは知見抽出対象語を数える際に重み付けに用いてもよい。すなわち、例えば番組名「ドラマA」に属する知見抽出対象語「登場人物X」を含むマイクロブログ記事が50件ある場合、確信度が1.0の場合は50件、確信度が0.8の場合は40件のマイクロブログ記事配信があったものとして件数を割り付けるようにしてもよい。または、確信度の低い評価表現の数が多くなった場合は、視聴者からの評価が下がっていると判断する知見としてもよい。

[0324] 以上説明した少なくとも1つの実施形態の知見抽出装置によれば、知見抽出対象語に関連付けられていない状態で随時配信される電子文書から知見抽出対象語に関する知見情報を随時抽出することが可能である。また、上記実施形態の少なくとも1つの知見抽出装置は、知見情報を最新の状態にすることができ。

[0325] なお、上記の各実施形態に記載した手法は、コンピュータに実行させることのできるプログラムとして、磁気ディスク（フロッピー（登録商標）ディスク、ハードディスクなど）、光ディスク（CD-ROM、DVDなど）、光磁気ディスク（MO）、半導体メモリなどの記憶媒体に格納して頒布することもできる。

[0326] また、この記憶媒体としては、プログラムを記憶でき、かつコンピュータが読み取り可能な記憶媒体であれば、その記憶形式は何れの形態であってもよい。

[0327] また、記憶媒体からコンピュータにインストールされたプログラムの指示に基づきコンピュータ上で稼働しているOS（オペレーティングシステム）や、データベース管理ソフト、ネットワークソフト等のMW（ミドルウェア）等が上記実施形態を実現するための各処理の一部を実行してもよい。

[0328] さらに、各実施形態における記憶媒体は、コンピュータと独立した媒体に

限らず、LANやインターネット等により伝送されたプログラムをダウンロードして記憶または一時記憶した記憶媒体も含まれる。

[0329] また、記憶媒体は1つに限らず、複数の媒体から上記の各実施形態における処理が実行される場合も本発明における記憶媒体に含まれ、媒体構成は何れの構成であってもよい。

[0330] なお、各実施形態におけるコンピュータは、記憶媒体に記憶されたプログラムに基づき、上記の各実施形態における各処理を実行するものであって、パソコン等の1つからなる装置、複数の装置がネットワーク接続されたシステム等の何れの構成であってもよい。

[0331] また、各実施形態におけるコンピュータとは、パソコンに限らず、情報処理機器に含まれる演算処理装置、マイコン等も含み、プログラムによって本発明の機能を実現することが可能な機器、装置を総称している。

[0332] なお、本発明のいくつかの実施形態を説明したが、これらの実施形態は、例として提示したものであり、発明の範囲を限定することは意図していない。これら新規な実施形態は、その他の様々な形態で実施されることが可能であり、発明の要旨を逸脱しない範囲で、種々の省略、置き換え、変更を行うことができる。これら実施形態やその変形は、発明の範囲や要旨に含まれるとともに、特許請求の範囲に記載された発明とその均等の範囲に含まれる。

符号の説明

[0333] 1, 100, 101, 102…知見抽出装置、2…ユーザ端末装置、10…制御部、11…記憶部、12…情報受信部、13…情報選択部、14…知見抽出部、15…情報分析部、16…入出力インタフェース、17, 26…バス、21…端末制御部、22…端末記憶部、23…表示部、24…入力部、25…通信インタフェース、111…第1手掛かり語辞書、112…知見抽出対象語辞書、113, 113a…第2手掛かり語辞書、141…知見抽出判定語チェック部、142…第1手掛かり語チェック部、143…第1概念抽出部、144…知見抽出対象語抽出部、145…第2概念抽出部、146…組み合わせ作成部、147…第1概念チェック部、151…知見抽出対

象語・第1概念・第2概念抽出部、152…電子文書割付部、153…知見抽出対象別判定部、18、19…知見評価更新装置、181…電子文書検索部、182…知見登録概念抽出部、183…評価部、184…確信度更新部、185…知見更新部

請求の範囲

- [請求項1] 電子文書から知見情報を抽出する知見抽出装置であって、
前記電子文書を受信する情報受信部と、
前記知見情報を抽出する対象語と、前記知見情報を抽出する手掛かり語とに基づいて前記電子文書から概念を抽出し、抽出された当該概念と前記対象語を関連付けた知見情報を作成する知見抽出部と、
前記抽出された知見情報を記憶する記憶部と、
前記知見情報が記憶された後、前記電子文書を前記記憶部内の知見情報に基づいて分析する情報分析部と、
を備える知見抽出装置。
- [請求項2] 前記電子文書に含まれる前記概念もしくは前記対象語を検索する電子文書検索部と、
前記検索部による検索の結果、前記電子文書に前記概念が規定数以上記載された場合に前記概念を抽出し、前記電子文書に前記対象語が規定数以上記載された場合に前記対象語に関連付けられた前記知見情報を前記記憶部から抽出する知見登録概念抽出部と、
前記知見登録概念抽出部が抽出した前記対象語もしくは前記概念に関連する数値情報の大きさもしくは変動を評価する評価部と、
前記数値情報の大きさもしくは変動の方向もしくは数値情報の変動の大きさに従って前記第概念と前記対象語の関連性の評価数値である確信度を更新する確信度更新部と、
前記確信度が規定値を下回った場合にユーザに判断を求め、ユーザの指示により選択中の前記概念間の関連を削除する知見更新部と、
を備える請求項1に記載の知見抽出装置。
- [請求項3] 前記手掛かり語はあらかじめ強度が設定され
前記知見抽出部は、
前記手掛かり語の強度と前記抽出された概念と前記対象語を関連付けた知見情報関連付けた知見情報を作成する請求項1及び請求項2の

いずれか 1 項に記載の概念間関連性評価更新システム。

[請求項4] 配信されたニュース情報から株取引に関する知見情報を抽出する知見抽出装置であって、
前記ニュース情報を受信する情報受信部と、
前記受信されたニュース情報から経済用語を含むニュース情報を選択する情報選択部と、
前記選択されたニュース情報から前記知見情報を抽出する知見抽出部と、
前記抽出された知見情報を記憶する記憶部と、
前記知見情報が記憶された後、前記情報受信部に受信された全てのニュース情報を前記記憶部内の知見情報に基づいて分析する情報分析部と
を備える知見抽出装置。

[請求項5] 請求項 4 に記載の知見抽出装置において、
前記知見抽出部は、
前記選択されたニュース情報が株価変動を評価する株価変動評価語を含むか否かを判定する株価変動評価語判定部と、
前記判定の結果、前記ニュース情報が株価変動評価語を含むとき、当該ニュース情報が銘柄のグループ名を定義するグループ定義語を含むか否かを判定するグループ定義語判定部と、
この判定の結果、前記ニュース情報がグループ定義語を含むとき、当該グループ定義語に基づいて当該ニュース情報からグループ名を抽出するグループ名抽出部と、
前記株価変動評価語判定部による判定の結果、前記ニュース情報が株価変動評価語を含むとき、当該ニュース情報から企業名及び証券コードを含む銘柄を抽出する銘柄抽出部と、
前記株価変動評価語判定部による判定の結果、前記ニュース情報が株価変動評価語を含むとき、当該ニュース情報から株価変動材料を抽出

する材料抽出部と、

前記抽出されたグループ名、銘柄及び株価変動材料を組み合わせる前記知見情報を作成する組み合わせ作成部と、

前記作成された知見情報を前記記憶部に書込む書込部と、
を備え、

前記知見情報は、前記グループ名、前記銘柄及び前記株価変動材料を含む情報、又は前記グループ名及び前記銘柄からなる第1情報と前記株価変動材料からなる第2情報とを含む情報である知見抽出装置。

[請求項6] 請求項5に記載の知見抽出装置において、

前記知見抽出部は、

前記株価変動材料の表現に対する手掛かり語と、前記株価変動の方向を表す情報と、前記株価変動の強度を表す情報とを関連付けて記憶した材料表現手掛かり辞書記憶部と、

前記抽出された株価変動材料に基づいて、前記材料表現手掛かり辞書を検索し、当該検索された株価変動の方向及び強度を表す各情報を当該株価変動材料に付与する部と

を更に備える知見抽出装置。

[請求項7] 請求項4に記載の知見抽出装置において、

前記情報受信部は、前記ニュース情報を受信する際に、当該ニュース情報にタイムスタンプを付加し、

前記知見抽出部は、前記選択されたニュース情報から所定の対象を示す表現を含むイベント情報を抽出すると共に、当該イベント情報が抽出されたニュース情報から抽出し、当該抽出したイベント情報及びタイムスタンプを前記知見情報として前記記憶部に書込む処理を実行し、

前記所定の対象を示す表現は、銘柄の企業名を表す企業名表現、又は前記銘柄のグループ名を定義するグループ定義語の直前の名詞を表す表現である知見抽出装置。

- [請求項8] 請求項4に記載の知見抽出装置において、
前記知見情報は上位概念の情報と下位概念の情報とによって構成される階層構造であり、
前記記憶部は評価情報と、前記下位概念の情報が前記上位概念の情報に帰属する度合いを示す確信度とを記憶し、
前記評価情報から抽出された、前記上位概念の情報と前記下位概念の情報とに関連する評価情報に基づいて評価値を算出する評価部と、
前記評価値に基づいて、前記確信度を更新する確信度更新部と、を備える知見抽出装置。
- [請求項9] 請求項5に記載の知見抽出装置において、
前記記憶部はさらに、評価情報と、前記知見情報に含まれる前記銘柄が前記グループ名に帰属する度合いを示す確信度とを記憶し、
前記ニュース情報に含まれている前記グループ名を確認する確認部と、
、
前記確認部により、前記ニュース情報に含まれていると確認されたグループ名に基づいて前記知見情報から銘柄を抽出する抽出部と、
前記銘柄に関連する前記評価情報に基づいて評価値を算出する評価部と、
前記評価値に基づいて、前記確信度を更新する確信度更新部と、を備える知見抽出装置。
- [請求項10] 請求項8に記載の知見抽出装置において、
前記確信度が規定値を下回った場合に、前記下位概念の情報と前記上位概念の情報とを含む知見情報を前記記憶部から削除する知見更新部を備える知見抽出装置。
- [請求項11] 記憶部を備え、配信された電子文書から概念を抽出する知見抽出装置に用いられるプログラムであって、
前記知見抽出システムを、
前記電子文書を受信する情報受信部、

前記知見情報を抽出する対象語と、前記知見情報を抽出する手掛かり語とに基づいて前記電子文書から概念を抽出し、抽出された当該概念を関連付けた知見情報を作成する知見抽出部、

前記抽出された知見情報を記憶する記憶部、

前記知見情報が記憶された後、前記電子文書を前記記憶部内の知見情報に基づいて分析する情報分析部、

として機能させるためのプログラム。

[請求項12] 記憶部を備え、配信されたニュース情報から株取引に関する知見情報を抽出する知見抽出装置に用いられるプログラムであって、
前記知見抽出装置を、

前記ニュース情報を受信する情報受信部、

前記受信されたニュース情報から経済用語を含むニュース情報を選択する情報選択部、

前記選択されたニュース情報から前記知見情報を抽出し、当該知見情報を前記記憶部に書込む知見抽出部、

前記知見情報の書込み後、前記情報受信部に受信された全てのニュース情報を前記記憶部内の知見情報に基づいて分析する情報分析部、

として機能させるためのプログラム。

[請求項13] 請求項12に記載のプログラムにおいて、
前記知見情報は上位概念の情報と下位概念の情報とによって構成される階層構造であり、

前記記憶部は評価情報と、前記下位概念の情報が前記上位概念の情報に帰属する度合いを示す確信度とを記憶し、

前記知見抽出装置を

前記評価情報から抽出された、前記上位概念の情報と前記下位概念の情報とに関連する評価情報に基づいて評価値を算出する評価部、

前記評価値に基づいて、前記確信度を更新する確信度更新部、

として機能させるためのプログラム。

- [請求項14] 上位概念の情報と下位概念の情報とによって構成される階層構造である知見情報と、前記知見情報の評価情報と、前記知見情報に含まれる前記下位概念の情報が前記上位概念の情報に帰属する度合いを示す確信度とを記憶する記憶部と
前記評価情報から抽出された、前記上位概念の情報と前記下位概念の情報とに関連する評価情報に基づいて評価値を算出する評価部と、
前記評価値に基づいて、前記確信度を更新する確信度更新部と、
を備える知見更新装置。
- [請求項15] 請求項14に記載の知見更新装置において、
前記記憶部は前記確信度の更新ルールを記憶し、
前記確信度更新部は、前記評価値と前記更新ルールとに基づいて前記確信度を更新するか否かを判定する知見更新装置。
- [請求項16] 請求項14に記載の知見更新装置において、
前記確信度に基づいて前記知見情報を更新する知見更新部を備える知見更新装置。
- [請求項17] 上位概念の情報と下位概念の情報とによって構成される階層構造の知見情報と、前記知見情報の評価情報と、前記知見情報に含まれる前記下位概念の情報が前記上位概念の情報に帰属する度合いを示す確信度とを記憶する記憶部を備え、前記知見情報を更新する知見更新装置に用いられるプログラムであって、
前記知見更新装置を、
前記評価情報から抽出された、前記上位概念の情報と前記下位概念の情報とに関連する評価情報に基づいて評価値を算出する評価部、
前記評価値に基づいて、前記確信度を更新する確信度更新部、
として機能させるためのプログラム。