

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7034184号
(P7034184)

(45)発行日 令和4年3月11日(2022.3.11)

(24)登録日 令和4年3月3日(2022.3.3)

(51)国際特許分類

F I
G 06 F 40/30 (2020.01) G 06 F 40/30
G 06 N 5/02 (2006.01) G 06 N 5/02

請求項の数 10 (全30頁)

(21)出願番号 特願2019-569743(P2019-569743)
 (86)(22)出願日 平成30年6月22日(2018.6.22)
 (65)公表番号 特表2020-525897(P2020-525897)
 A)
 (43)公表日 令和2年8月27日(2020.8.27)
 (86)国際出願番号 PCT/IB2018/054610
 (87)国際公開番号 WO2019/003069
 (87)国際公開日 平成31年1月3日(2019.1.3)
 審査請求日 令和2年11月30日(2020.11.30)
 (31)優先権主張番号 15/632,564
 (32)優先日 平成29年6月26日(2017.6.26)
 (33)優先権主張国・地域又は機関
 米国(US)
 (31)優先権主張番号 15/844,723
 (32)優先日 平成29年12月18日(2017.12.18)
 最終頁に続く

(73)特許権者 3900009531
 インターナショナル・ビジネス・マシ
 ンズ・コーポレーション
 INTERNATIONAL BUSI
 NESS MACHINES CORPO
 RATION
 アメリカ合衆国10504 ニューヨー
 ク州 アーモンク ニュー オーチャード
 ロード
 New Orchard Road, A
 rmonk, New York 105
 04, United States of
 America
 (74)代理人 100108501
 弁理士 上野 剛史
 最終頁に続く

(54)【発明の名称】 意味グラフにおけるメタ関係の適応評価のためのコンピュータ実施方法、コンピュータ・システム、及びコンピュータ・プログラム

(57)【特許請求の範囲】

【請求項1】

コンピュータの情報処理により、意味グラフにおけるメタ関係の測定を行う方法であって、ノードの形態の概念がエッジの形態の意味関係によりリンクされる、知識ベースに基づいた自然言語ソースから前記意味グラフを得ることと、前記意味グラフのエッジ及びノードのメタデータに、メタ関係を測定するための特徴ベクトルとしての重み付けを符号化(エンコード)することであって、前記メタ関係は、前記意味グラフの概念に適用され、かつ、前記意味グラフのエッジによって定義される前記意味関係とは独立しており、前記特徴ベクトルは、前記メタ関係の強度を含み、前記強度は前記ノードの実行時入力を使用する、符号化することと、

意味論的グラフの1又は複数の概念に関連する入力文脈を用いてグラフを活性化することであって、前記重み付けが意味論的グラフを介して広がっていく活性化信号に適用され、意味論的グラフの概念のサブセットに対する前記メタ関係の測定値を生成する、活性化することと、

を含む、方法。

【請求項2】

前記入力文脈に対してグラフ活性化を実行することは、前記入力文脈における概念エンティティを検出すること、並びに、前記意味グラフにおける前記概念に対応するノードを活性化させること、信号に前記重み付けを適用しながら前記信号を隣接するノードに向けて順次伝達して前記隣接するノードを活性化させること、及び結果としての再活性化信号を

有する 1 つ又は複数の焦点ノードを特定することと、
を含む、請求項 1 に記載の方法。

【請求項 3】

前記入力文脈における前記メタ関係の測定を反映した、前記意味グラフの結果的に活性化された部分を出力することをさらに含む、請求項 2 に記載の方法。

【請求項 4】

前記意味グラフが基づく知識ベースとは独立したリソースのセットから得られたメタ関係の測定値に対する重み付けを前記グラフに適用することをさらに含む、請求項 1 乃至 3 のいずれかに記載の方法。

【請求項 5】

前記重み付けは、前記リソースのセットから得られる未加工値であり、前記リソースのセットへの追加に応答して更新され、前記未加工値は、1 つのグラフ活性化の処理の中で適用される、請求項 4 に記載の方法。

【請求項 6】

前記重み付けは、前記メタ関係の異なる観点に対する多次元測定、及び / 又は前記メタ関係の極性を示す、請求項 1 乃至 5 のいずれかに記載の方法。

【請求項 7】

前記重み付けは、入力文脈の概念エンティティに対するノードの実行時入力に応じて計算される特徴ベクトルである、請求項 1 乃至 6 のいずれかに記載の方法。

【請求項 8】

前記メタ関係は、

感情分析、バイアス評価、予測分析におけるバイアス、情報検索を用いたクエリ拡張、リスク評価、地理空間推定、及び臨床試験マッチングを含む治療、使用又は処理の適合性を含む群のうちの 1 つの形態の現象に関する、請求項 1 乃至 7 のいずれかに記載の方法。

【請求項 9】

請求項 1 乃至 8 のいずれかに記載の方法を、プロセッサとメモリを有するコンピュータ・ハードウェアとして構成した、コンピュータ・システム。

【請求項 10】

請求項 1 乃至 8 のいずれかに記載の方法を、コンピュータに実行させる、コンピュータ・プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、意味グラフに関し、より具体的には、意味グラフにおけるメタ関係の適応評価に関する。

【背景技術】

【0002】

構造化された知識から派生する意味概念及び関係のグラフは、高精度の自然言語処理 (NLP) に対する極めて有用なリソースである。

【0003】

オントロジ (存在論) は、コンピュータサイエンスにおいて、「共有概念化の明示的仕様」 (explicit specification of a shared conceptualization) と定義され、概念化は、所与のタスク要件に対する要件について、実世界の意味論 (semantics) の何らかのサブセットであり得る。それは、オブジェクトの概念又はクラス、オブジェクト特性、及び概念間関係、並びにターゲットドメインにおけるこれらのインスタンスを含有し得る。こうした構造化リソースは、ドメイン知識の共有及び再利用を促進し、かつ、NLP アプリケーションには貴重である。こうしたリソースの主な例は、米国国立医学図書館 (NLM 、 2013) により提供された統一化医学言語体系 (Unified Medical Language System, UMLS) である。データセットは、医学ドメインにおける概念及び概念間関係のオントロジと関連して、膨大なインスタンス表層形式など、大量の語彙 (lexicon) からな

10

20

30

40

50

る。

【0004】

活性化拡散 (spreading activation) 理論は、ヒトの脳が自然言語の意味を処理し、理解できる手法を人工的に表すために、ヒトの意味記憶の抽象モデルにおいて最初に提案された。このモデルは、検索タスクに対して強化され、認識心理学から神経科学 (neuroscience) に至るまで、特に自然言語処理など、他の多くの関連分野における研究に対したインスピレーションを与えた。

【0005】

活性化拡散の根本的な前提は、人工知能における結合説のものに関連し、ニューラル・ネットワークに対して類似のモデルを用いてヒトの脳における電気信号の扇形展開を反映する。ニューラル・ネットワークの場合、グラフの頂点は、単一のニューロンを表し、エッジは、シナプスを表すことができる。情報検索及び語義の曖昧性解消 (word-sense disambiguation) において、一般的に、頂点は語意を、エッジは、語意間の、語彙又は意味のリンクのいずれかの、何らかの形式の関係を表す。

10

【0006】

活性化拡散戦略は、非特許文献 1 に記載される。この技術は、テキストの文書を処理し、ノードからの信号を伝搬させ、次にグラフを通じて拡散させることにより、ノードが見つかると意味グラフ内のノードを活性化させることを含み、テキスト内になかった他のノードに潜在的に蓄積する。

【0007】

この基本的な考え方は多様に実装されているが、IBM (登録商標) の Galaxy (IBM (登録商標) 及び Galaxy は、インターナショナル・ビジネス・マシーンズ・コーポレーションの商標である) により実装されるものは、語義曖昧性解消及び語義推定 (word-sense inference) に有用である。

20

【0008】

NLP には、エンティティ間の関係に関する現象を分析する様々な分野がある。こうした分野の例として、感情分析、バイアス検出、地理空間推定 (geo-spatial inference)、文脈関連、及びリスク評価 (risk assessment)、エンティティ間のこうした現象の正確なスコア化は、従来技術の方法において必要であり、こうした現象の分析の改善が必要である。

30

【先行技術文献】

【非特許文献】

【0009】

【文献】Ronan Mac an t Saoir 著、「Using Spreading Activation to Evaluate and Improve Ontologies」、COLING 2014、第 25 回 International Conference on Computation Linguistics の予稿集:技術論文、2237 ~ 2248 ページ、ダブリン、アイルランド、2014 年 8 月 23 ~ 29 日

【発明の概要】

【発明が解決しようとする課題】

40

【0010】

意味グラフにおけるメタ関係の適応評価のためのコンピュータ実施方法、コンピュータ・システム、及びコンピュータ・プログラムを提供する。

【課題を解決するための手段】

【0011】

本発明の 1 つの態様によれば、意味グラフにおけるメタ関係の適応評価のためのコンピュータ実施方法であって、グラフ・ノードの形態の概念がグラフ・エッジの形態の意味関係によりリンクされる、知識ベースに基づいた意味グラフを提供することと; 意味グラフのエッジ及びノードのメタデータにおいて、メタ関係を測定するための重み付けを符号化することであって、メタ関係は、意味グラフの概念に適用され、かつ、意味グラフのエッジ

50

により定義される意味関係から独立している、符号化することと；意味グラフの1つ以上の概念に関連する入力文脈に対してグラフ活性化を実行することであって、重み付けは、意味グラフを通して活性化拡散信号に適用されて意味グラフの概念のサブセットに対するメタ関係の尺度を生成する、実行することと、を含む方法が提供される。

【0012】

記載される方法は、グラフ活性化を用いて意味グラフのサブセットに定義される概念間のメタ関係をスコア化する利点をもたらす。メタ関係に対する符号化された重み付けは、入力文脈の概念に対して活性化され、相関関係にあるエンティティのメタ関係の複雑なスコア化及び分析が可能になり得る。

【0013】

入力文脈に対してグラフ活性化を実行することは、入力文脈における概念のインスタンスを発見し、意味グラフにおける概念に対応するノードを活性化させることと；信号に重み付けを適用しながら信号を隣接するノードに向けて順次伝達して隣接するノードを活性化させることと；結果としての最活性化信号を有する1又は複数の焦点ノードを特定することと、を含むことができる。本方法は、入力文脈におけるメタ関係の測定を反映する意味グラフの結果としての活性化部分を出力することができる。このことは、入力文脈の概念のインスタンス間のメタ関係をスコア化する利点をもたらす。

10

【0014】

本方法は、意味グラフが基づく知識ベースから独立したリソースのセットから得られたメタ関係の尺度に対する重み付けでグラフをシードすることができる。これにより、特定のアプリケーションに対する重み付けの符号化が可能になる。意味グラフは、メタ関係の尺度を重ね合わせることができるベースとして用いることができる。

20

【0015】

重み付けは、メタ関係及び／又はメタ関係の極性の異なる観点 (aspect) に対する多次元測定を示し得る。メタ関係は、重み付けにおける次元として測定し符号化することができる様々な観点を有し得る。重み付けは、ポジティブ及びネガティブを含み得る極性も有することができる。

【0016】

1つの実施形態において、重み付けは、リソースのセットから得られる未加工値とし、リソースのセットへの追加に応答して更新することができ、未加工値は、グラフ活性化の際に適用される。

30

【0017】

別の実施形態において、重み付けは、入力文脈の概念のインスタンスに対するノードに対する実行時入力に応じて計算できる特徴ベクトルとすることができます。これにより、スコア化される入力文脈に対する入力に基づく、重み付けを実行時に適用する利点が可能になる。特徴ベクトルは、ノードに対する実行時入力に適用される関連係数 (relevance factor) を含むことができ、関連係数は、異なるノードによって異なり得る。

【0018】

さらに、特徴ベクトルは、メタデータ関係特徴に加えて、入力文脈における概念のインスタンスに対する意味又は語彙特徴を含み得る。これにより、意味グラフの意味関係を分析に組み込むことが可能になる。代替的に、特徴ベクトルは、メタ関係にのみ関連し得る。

40

【0019】

特徴ベクトルは、統計技術を用いて重み付けの信頼度スコア及びメタ関係の集約を定義することができる。

【0020】

メタ関係は、感情分析、バイアス評価、予測分析におけるバイアス、情報検索を用いたクエリ拡張、リスク評価、地理空間推定、及び臨床試験マッチングを含む治療、使用又は処理の適合性の群のうちの1つの形態の現象に関連し得る。これらの現象は、意味グラフの概念にわたり普遍的に適用することができる。

【0021】

50

本発明の別の態様によると、意味グラフにおけるメタ関係の適応評価のためのシステムであって、プロセッサ、及びプロセッサに、コンポーネントの機能を実行させるためのコンピュータ・プログラム命令を提供するように構成されたメモリと；グラフ・ノードの形態の概念がグラフ・エッジの形態の意味関係によりリンクされる、知識ベースに基づいた意味グラフを提供するための意味グラフ・コンポーネントと、意味グラフのエッジ及びノードのメタデータにおいて、メタ関係を測定するための重み付けを符号化するためのメタ関係コンポーネントであって、メタ関係は、意味グラフの概念に適用されかつ意味グラフのエッジにより定義される意味関係から独立している、メタ関係コンポーネントと、意味グラフの1又は複数の概念に関連する入力文脈に対してグラフ活性化を実行するための実行時コンポーネントと、を含み、重み付けは、意味グラフを通して活性化拡散信号に適用されて意味グラフの概念のサブセットに対するメタ関係の尺度を生成する。

【 0 0 2 2 】

システムは、入力文脈の概念に対するメタ関係のスコア化が複雑な場合、入力文脈に自動的に適用されるメタ関係重み付けの符号化を可能にする利点をもたらし得る。

【 0 0 2 3 】

実行時コンポーネントは、入力文脈における概念のインスタンスを発見するための概念検知コンポーネントと、意味グラフの概念に対応するノードを活性化させ、信号に重み付けを適用しながら信号を隣接するノードに向けて順次伝達して隣接するノードを活性化させ、結果としての最活性化信号を有する1又は複数の焦点ノードを特定するためのグラフ活性化コンポーネントと、を含むことができる。

【 0 0 2 4 】

システムは、入力文脈におけるメタ関係の測定を反映する意味グラフの結果としての活性化部分を出力するための出力コンポーネントを含むことができる。出力コンポーネントは、ノード及びエッジに対する活性化重み付けがメタ関係の測定を表す活性化サブグラフを出力するための環境文脈コンポーネントを含み得る。

【 0 0 2 5 】

システムは、意味グラフが基づく知識ベースから独立したリソースのセットから得られたメタ関係の尺度に対する重み付けでグラフをシードするための重み付け追加コンポーネントを含み得る。システムは、さらに、リソースのセット又は入力文脈の追加に応じて重み付けを更新するための重み付け更新コンポーネントを含み得る。

【 0 0 2 6 】

システムは、感情分析、バイアス評価、予測分析におけるバイアス、情報検索を用いたクエリ拡張、リスク評価、地理空間推定、及び臨床試験マッチングの群のうちの1つの形態の治療、使用又は処理の形態の現象に関するメタ関係を測定するためのシステムに組み込むことができる。

【 0 0 2 7 】

本発明のさらに別の態様によると、意味グラフにおけるメタ関係の適応評価のためのコンピュータ・プログラム製品が提供され、コンピュータ・プログラム製品は、プログラム命令がそこに具体化されたコンピュータ可読ストレージを含み、プログラム命令は、プロセッサにより実行可能脳であり、プロセッサに、グラフ・ノードの形態の概念がグラフ・エッジの形態の意味関係によりリンクされる、知識ベースに基づいた意味グラフを提供させることと；意味グラフのエッジ及びノードのメタデータにおいて、メタ関係を測定するための重み付けを符号化することであって、メタ関係は、意味グラフの概念に適用されかつ意味グラフのエッジにより定義される意味関係から独立している、符号化することと；意味グラフの1又は複数の概念に関連する入力文脈に対してグラフ活性化を実行することであって、重み付けは、意味グラフを通して活性化拡散信号に適用されて意味グラフの概念のサブセットに対するメタ関係の尺度を生成する、実行することと、を行わせる。

【 0 0 2 8 】

本発明とみなされる主題は、本明細書の結論部にある特許請求の範囲において具体的に指示され、明確に特許請求される。本発明は、構成及び動作の方法の両方に関して、その

目的、特徴及び利点とともに、以下の詳細な説明を添付図面と共に読むときに参照することによって、最も良く理解することができる。

【0029】

ここで、本発明の好ましい実施形態が、単に例示の目的で、以下の図面を参照して説明される。

【図面の簡単な説明】

【0030】

【図1】本発明による方法の態様の例示的実施形態の流れ図である。

【図2】本発明による方法のさらに別の態様の例示的実施形態の流れ図である。

【図3】本発明によるシステムの例示的実施形態のブロック図である。

10

【図4】本発明による方法の第1の例示的実施形態を示す概略図である。

【図5】本発明による方法の第2の例示的実施形態を示す概略図である。

【図6】本発明による方法の第2の例示的実施形態を示す概念図である。

【図7】本発明が実装され得るコンピュータ・システム又はクラウド・サーバの実施形態のブロック図である。

【図8】本発明が実装され得るクラウド・コンピューティング環境の概略図である。

【図9】本発明が実装され得るクラウド・コンピューティング環境の抽象化モデル層の図である。

【発明を実施するための形態】

【0031】

例示を簡明かつ明確にするために、図に示される要素は、必ずしも縮尺通りに描かれていないことを認識されたい。例えば、要素のうちの幾つかの寸法は、明確にするために、他の要素に対して誇張されていることがある。さらに、適切と考えられる場合、参照符号は、対応する又は類似する特徴を示すために、図の間で繰り返されることがある。

20

【0032】

記載される方法は、自然言語処理の際、意味グラフにおける活性化拡散を用いて、ドメインにおけるエンティティ間の現象の「メタ関係 (meta-relationship)」の強度を定量化（測定）する。メタ関係、及び拡散信号がこの現象の強度の尺度（測定値）をどのように生成するかの詳細は、グラフ・メタデータに符号化され、タスクの性質及び当該ドメインに応じて変化する。

30

【0033】

メタ関係は、グラフのエンティティに普遍的に適用される属性又は特性の形態の現象を説明する。現象の強度は、意味グラフにシードしておき、追加の入力ソースが利用できる際に、活性化拡散などのグラフ活性化を用いて適応測定することができる。

【0034】

これは、ドメインに対する入力文脈におけるエンティティ間の現象をスコア化する問題に対する技術的解決法を与える。本解決法は、エンティティ間の複雑なスコア化及び測定を可能にする意味グラフのグラフ活性化を用いる。

【0035】

「メタ関係」という用語は、エンティティに適用される現象に関連するメタレベル関係又は普遍的関係を説明するのに用いられる。「メタ」の解釈は、より高レベルのものであります既存の意味関係グラフからは独立しているものである。メタ関係は、いずれかの特定のグラフの既存の意味上に重ね合わされる。

40

【0036】

メタ関係は、初期意味グラフの一部ではなく、必ずしもグラフのエッジにより表されるエンティティ間の関係に関連しない。意味グラフは、「著者」及び「記事」等の2つのノード又は概念間の「著作 (w r o t e)」等の意味関連に対するエッジを含有する一方、記載される方法は、この特定の意味的関係の性質又は重要性に焦点を当てない。代わりに、この構造を利用して、グラフのエンティティに全体的又は普遍的に、しかし異なる強度で適用することができる「感情」、「バイアス」、「地理空間関連」又は「リスク危険度 (r

50

isk exposure)」等の現象、属性、又は特徴のより高レベルのメタ文脈を定量化する。

【0037】

伝統的意味グラフ用語は、自然言語処理における意味グラフの主要機能である、概念間の意味関係に対するグラフ・ノード及びエッジを参照する。記載される方法は、測定され得る追加の観点又は現象の重み付けを重ね合わせる。この追加の現象は、ノードの個別のリンクされた対の詳細に限定されない。

【0038】

意味グラフは、自然言語処理の間、メタ関係の対象の分析を実施するために、初期意味グラフを提供するために用いられたメタ関係に対する追加リソースを用いて、文脈からメタ関係の強度の尺度でシードされる。

10

【0039】

メタ関係は、グラフに格納された情報を処理することにより、露出され、強度が割り当てられる。意味グラフにおける既存の関係に追加したメタ関係の拡張は、意味グラフがベースとするソースの第1のデータベースに加えて追加のリソースの第2のデータベースを必要とする。

【0040】

意味グラフのデフォルト解釈は、同じ意味グラフを自然言語処理においてどのように用い得るかを本質的に変更するメタ関係により表される任意の数の様々な普遍的現象を用いる記載される方法で増強される。ユーザが、通常いずれかの意味グラフを通して信号を送ると、ユーザは、重み付けにより、ものごとが意味的にどれだけ密接に関連しているかが示されると期待する。これは、文脈においていかなるメタ関係が好まれる何らかのことを示す。これが「ものごとが地理的にどれだけ近いか」又は「ものごとが互いに対してどれだけ対してどれだけバイアスされているか」等であり得る。

20

【0041】

活性化拡散重みを用いて意味関係の強度を定量化することは、既に存在する。記載される方法は、これを拡張して、メタ関係強度を用いる単純文脈関連以外の何らかの現象を定量化する。

【0042】

意味グラフは、特定のノード及びエッジ形式に重み修飾子を起因させるか、又は特定のノードに対する開始重みを修正することにより、ノード及びエッジ・レベルでカスタマイズすることができる。記載される方法は、より高レベルの文脈に関してグラフ構造及びコンテンツのドメイン・カスタマイズ済み活性化拡散を提供する。コーパス関連、バイアス、又はリスク危険度等の現象又は文脈のメタ関係の認識は、自然言語処理に対して非常に異なる考慮事項及び結果を有している。

30

【0043】

特定の言語現象の強度の程度の検出又は分析の何らかの対象の特定の特徴を定量化することに焦点を当てる自然言語処理のドメインには多くのトピックがある。記載される方法は、同じ量又は現象の環境表現を送出するために、多くのこうしたタスクに対して適用できる技術を提供する。バイアス又は地理関連等の現象の環境表現は、本質的には、ノード及びエッジの活性化重みが現象の強度の程度を表す活性化サブグラフである。ここで環境という語は、表現のネットワーク又はウェブ様構造を反映する。

40

【0044】

自然言語処理の以下の特定の分野：感情分析、バイアス検出、地理空間推定、文脈関連及びリスク評価等における例示的実施形態を説明する。

【0045】

図1を参照すると、流れ図100は、グラフ活性化を用いる適応評価のために、現象に対するメタ関係を意味グラフに適用する記載される方法の態様の例示的実施形態を示す。

【0046】

知識ベースを提供する意味ソース又はリソース110に対して、構築すること又は既存のグラフとすることができる意味グラフ111が提供される(101)。これに対するリソ

50

ースは、解析フレーム、オントロジ、固有表現抽出（named entity recognition、NER）データの共起等であり得る。この記載される方法において、意味グラフ111は、記載される方法により用いることができる既存の予め構築されたグラフとすることができる、その場合、オリジナルのリソース110を再訪する必要はない。

【0047】

本方法は、自然言語テキスト又はドメイン用語の任意の特定のスタイルに特有ではない。本明細書で用いられる意味グラフは、固有表現抽出（NER）及び自然言語構文解析プログラムを適用できる任意の自然言語ソースから誘導できる。

【0048】

本方法は、関心ある現象について参照する追加リソース112の形態のインスタンスのコードパスを識別することができる（102）。メタ関係を用いて強度を測定する特定の現象に関して、グラフ活性化に対する初期メタデータのベースとする追加リソース112の形態の例が用いられる。追加リソース112は、関心ある現象の文脈において意味グラフ111のエンティティを参照する。

10

【0049】

本方法は、シードとして用いることができるNLPタスクの目的である現象を反映するエンティティの全ての対の間の既存の強度又は極性のメタ関係に対する重み付けを導出することができる（103）。これは、感情分析又はバイアス検出データ、地理的距離データ、IR統計、リスク係数スコア等の目標の従来の分析を用いて提供され得る。本方法は、背景知識を利用して、タスク及びドメインに従って意味グラフを重み付けする。

20

【0050】

本方法は、メタデータにおいて、メタ関係重み付け113をハードコード化未加工値又は特徴ベクトルとして符号化することができる（104）。ノード及びエッジ重み付けは、入力データ特徴の関数であり、後の計算用に未加工値又は特徴ベクトルとして格納することができる。

【0051】

従来の意味グラフ実装用のメタデータ値又は特徴ベクトルは、通常、グラフのノード又はエッジにおける拡散信号の蓄積を反映する単一の正の値である。記載される方法において、与えられる重み付けは、評価されるメタ関係に関連する新しい意味を有する。

30

【0052】

特徴ベクトルは、何らかのオブジェクトを表す数値的特徴のn次元ベクトルである。特徴ベクトルは、統計要領で用いられる説明変数のベクトルと等価である。特徴ベクトルは、予測を行うためにスコアを決定するのに用いられる線形予測関数（linear predictor function）を構成するためにドット積を用いる重み付けと組み合わされることが多い。

【0053】

特徴ベクトルは、以下を含み得るが、これらに限定されない。

i . コードパスの概念のインスタンスに関する意味及び語彙特徴、

i i . タスクの主要目的であるメタ関数の強度。これは、例えば、分極できるバイアス又は感情の場合のように、多次元のものとすることができます、

i i i . 上記の全てに関する信頼度スコア、及び

i v . 同一ベクトルの論理回帰等の統計技術を用いた関連及び強度の尺度の集約。

40

【0054】

特徴ベクトルにおけるメタ関係の強度には、ハードコード化関連係数と組み合わせができるノードについての実行時入力を用い得る。これは、各ノード又はノードの型式について異なり得る。

【0055】

ノード間の物理リンクという意味における関係は、修正されない。代わりに、メタ関係は、活性化重み付けの数値及び値の符号（+ / -）に反映される。メタ関係のラベル（バイアス又はリスク危険度）は、ノード又はエッジのグラフでは符号化されない。変化は、活性化拡散後の信号強度が新しい意味を有することである。

50

【 0 0 5 6 】

実行時に用いられる未加工値又は特徴ベクトルを保存すること（105）により、本方法は、システムへの将来の入力の性質に従ったグラフ重み付けの随意的適応も容易にする。

【 0 0 5 7 】

本方法は、感情分析、意見バイアス検出、地理空間関連、意味関連、リスク危険度などの意図に関係なく、有意の適応を必要としない。

【 0 0 5 8 】

図2を参照すると、流れ図150は、グラフ活性化を用いて現象に関するメタ関係の適応評価の記載される方法の別の態様の例示的実施形態を示す。

【 0 0 5 9 】

実行時に入力文脈114の形態の初見のテキストが処理されると（151）、方法は、エンティティ基準を発見し（152）、それらの開始点において意味グラフ115を活性化させることができる。通常、これは、意味グラフを構築する（101）際に用いるのと同じNER生成器を用いることができる。

【 0 0 6 0 】

本方法は、図1の方法で追加されたノード及びエッジにおけるメタ関係の重み付けに従って信号を伝搬することにより、意味グラフにおける活性化拡散を実行することができる（153）。

【 0 0 6 1 】

インスタンスについて一意の識別子のセットを用いて、入力文脈114における概念のインスタンスを発見することにより、対応するノードが意味グラフにおいて活性化され、信号は、そこから隣接するノードに向かって伝達し、隣接するノードを活性化する。

【 0 0 6 2 】

1つの実施形態において、入力文脈で発見された概念のインスタンスは、ノードの重み付けのハードコード化未加工値を用いて対応するノードを活性化させることができる。別の実施形態において、概念のインスタンスからの入力は、ノードの特徴ベクトルの計算に用いることができる。入力は、重み付け特徴ベクトルに定義された通り異なるノード又はノードの形式に対して異なるハードコード化関連係数と組み合わせることができる。

【 0 0 6 3 】

信号がソース・ノードから拡散して離れると、この信号は、グラフのノード及びエッジに対して関連付けられた重み付けモデルで指定された量だけ弱くなる。この信号が複数の近隣ソース・ノードから拡散する場合、信号は組み合わせられ、重なり点は、より大きい程度まで活性化される。最も活性化を蓄積するノードは、入力文脈に対する焦点ノードとみなされる。その結果のグラフの活性化部分は、入力文脈のメタ関係の現象の固有の強度を反映する。

【 0 0 6 4 】

入力文脈114は、活性化終了時の信号に従った検出されたエンティティにより、スコア化し（154）、このスコアを、結果として出力することができる。

【 0 0 6 5 】

スコアは、環境文脈ノードに対して用いて、同一文脈における更に別の結果を導出することができ（155）、それを更に別の結果として出力することができる。例えば、入力として提供された特定の法律に対する関連地理的管轄権の推定である。

【 0 0 6 6 】

記載される方法は、分析される文脈におけるエンティティの明示的発生に依存せず、グラフにおける環境接続を介して追加の外部概念の関連付け及びスコア化が可能である。本方法は、間接的に関連するエンティティの環境文脈により、特定の文脈における関係の強度の評価に影響を与えることを可能にし、より包括的な見地をもたらす。

【 0 0 6 7 】

本方法は、文脈全体の単に集約された「要点」よりも、むしろ詳細なエンティティ・レベル関係及びエンティティの対の間の関係強度のスコア化を容易にする。

10

20

30

40

50

【0068】

幾つかの実施形態において、入力文脈 114 が処理された後に信号がどのように変化したかに従ってメタ関係重み付け 116 を更新すること (156) が有用であり得る。これには、図 1 に用いられる重み付けを導出する (103) 方法を用いることができる。例えば、著者とリストに載っていないトピック間の歴史的バイアスは、入力文脈 114 を用いた活性化の結果と異なる場合がある。

【0069】

図 3 を参照すると、ブロック図により、意味グラフ評価システム 210 の形態の記載したシステムの例示的実施形態としてシステム 200 を示す。

【0070】

意味グラフ評価システム 210 は、少なくとも 1 つのプロセッサ 201、ハードウェア・モジュール、又は少なくとも 1 つのプロセッサ上で実行されるソフトウェア・ユニットとすることができる記載したコンポーネントの機能を実行するための回路を含む。並行処理スレッドを実行する複数のプロセッサを設けて、コンポーネントの機能の一部又は全部の並行処理を行わせることができる。メモリ 202 は、少なくとも 1 つのプロセッサ 201 にコンピュータ命令 203 を与えて、コンポーネントの機能を実行させるように構成することができる。

【0071】

意味グラフ評価システム 210 は、意味リソース 110 に基づいて評価される現象に対するメタ関係重み付け 113 を意味グラフ 111 に追加することにより、意味グラフ作成システム (図示せず) の一部として、又は意味グラフにおけるメタ関係の適応評価用の独立したシステムとして提供することができる。意味グラフ評価システム 210 は、メタ関係に関する重み付け値をシードするため評価される現象に関連する追加リソース 112 を用いることができる。入力文脈 114 は、グラフ活性化を用いて現象を評価するために実行時に提供することができる。

【0072】

意味グラフ評価システム 210 は、グラフ・ノードの形態の概念がグラフ・エッジの形態の意味関係によりリンクされる知識ベースを形成する意味リソース 110 に基づいて意味グラフを提供するための意味グラフ・コンポーネント 211 を含むことができる。

【0073】

意味グラフ評価システム 210 は、意味グラフ 111 のエッジ及びノードのメタデータにおいて、メタ関係の強度に対する重み付けを符号化するためのメタ関係コンポーネント 220 を含むことができる。メタ関係は、意味グラフ 111 の概念に適用することができ、意味グラフのエッジにより定義される意味関係からは独立し得る。

【0074】

メタ関係コンポーネント 220 は、メタ関係重み付けがベースとし得る追加リソース 112 を選択するための追加リソース・コンポーネント 221 を含むことができる。また、メタ関係コンポーネント 220 は、追加リソース 112 から得られたメタ関係に対して使用するシード重み付けを導出するためのシード重み付け導出コンポーネント 222 を含むこともできる。メタ関係コンポーネント 220 は、意味グラフ 111 がベースとする知識ベースから独立した追加リソース 112 のセットから得られたメタ関係の強度の重み付けで意味グラフをシードするための重み付け追加コンポーネント 223 を含むこともできる。

【0075】

メタ関係コンポーネント 220 は、追加リソース 112 又は入力文脈 114 への追加に応答して重み付けを更新するための重み付け更新コンポーネント 224 を含むことができる。

【0076】

意味グラフ評価システム 210 は、意味グラフの 1 又は複数の概念に関連する入力文脈 114 に対するグラフ活性化を実行するための実行時コンポーネント 230 を含むことができる。重み付けを適用して意味グラフ 111 を通して活性化信号を拡散し、意味グラフの概念のサブセットに対するメタ関係の強度の尺度を生成することができる。

10

20

30

40

50

【0077】

実行時コンポーネント230は、評価する入力文脈114を選択するための入力文脈選択コンポーネント231を含むことができる。実行時コンポーネント230は、入力文脈114において概念のインスタンスを発見するための概念検出コンポーネント232を含むことができる。

【0078】

実行時コンポーネント230は、意味グラフ111における概念に対応するノードを活性化させ、信号に重み付けを適用しながら信号を隣接ノードに向けて順次伝達して隣接ノードを活性化させ、結果としての再活性化信号を有する1又は複数の焦点ノードを決定するための、メタ関係のためのグラフ活性化コンポーネント233を含むことができる。

10

【0079】

実行時コンポーネント230は、グラフ活性化に対応してノードをスコア化するためのスコア化コンポーネント234と、ノード及びエッジに対する活性化重み付けがメタ関係の強度の程度を表す活性化サブグラフのための環境文脈コンポーネント235とを含むことができる。

【0080】

実行時コンポーネント230は、入力文脈におけるメタ関係の強度を反映する意味グラフの結果としての活性化部分を出力するための出力コンポーネント237を含むことができる。

【0081】

20

メタ関係を評価する異なる現象の様々な例を、以下の項に説明する。

【0082】

臨床試験マッチング

例示的実施形態において、意味グラフは、「腫瘍」、「プロゲステロン」、又は「リンパ節」などの医学ドメイン概念について統合医学用語システム(UMLS)データから構築するか、又は既存のグラフを用いることができる。

【0083】

この意味グラフと共に、成功した治験候補者の注釈付き患者記録のコーパスを用いて、このタスクのグラフ活性化の重み付けを構成することができる。代替的に、医学専門家から得た、試験における許容基準についてのハードコード値のセット又は値の範囲を同じ様式で用いることもできる。

30

【0084】

ノード重み付けは、特徴ベクトルとして、臨床試験に対するノードの関連性に対して構成することができる。特徴ベクトルは、ノードへの特定の患者の入力を入力として用いる。

【0085】

本例では、最終スコアは、患者が臨床試験の要件に合致する程度を符号化する。グラフのノード及びエッジは、「腫瘍」、「プロゲステロン」、又は「リンパ節」などの医学ドメイン概念とすることができます。この概念の各々は、試験される薬に応じて異なる関連性を有す、合致の強度は、この概念に対して評価される患者が示す値と併せて従来のハードコード関連性の関数である。

40

【0086】

従って、特徴ベクトルは、ノードごとに異なり得る。例の幾つかを以下に挙げる。

a) 「Tumour Size(腫瘍サイズ)」ノードは、ハードコード関連係数0.9を有するので、患者が到着し、この概念に対する関連値を有すると、ベクトルは、(患者の腫瘍サイズ*0.9)の正規化値を含み得る;

b) 「Progesterone Level(プロゲステロン・レベル)」ノードが、この薬に対して0.8の関連性を有する場合、ベクトルは、(患者のプロゲステロン・レベル*0.8)を含む;

c) 「ki67増殖指数」に関するノードは、正規化百分率(n%*0.7)を有し得る。

【0087】

50

次に、患者の使用可能な値は、使用可能でない特徴についてのデフォルト値を用いて、グラフを活性化するために用いられ、重要な概念の各々における信号の正味蓄積は、この患者の、臨床試験の総合的適合性を反映する。

【0088】

感情分析

感情分析は、対象の意見を抽出し、それをポジティブ (positive、肯定的)、ネガティブ (negative、否定的)、もしくはニュートラル (neutral、中立) として分類し、又はこの能力を拡張して何らかの数値表現を行うことを含む。通常、感情分析は文書 (document) レベルで適用され、文書の集約された極性が議論のトピックに適用される。語句レベルの感情検出は、「観点レベル感情分析 (aspect level sentiment analysis)」として知られるトピックの特定の要素にますます適用されている。この場合、トピック全体の各サブ要素について、別個のスコアがドキュメント又はコーパスに集約される。本プロセスの最も困難な部分の1つは、語句「ネガティブな試験結果」対「ネガティブな経験」など、文脈において語の固有の感情がいつ変化するかを認識することである。

10

【0089】

感情分析は、記載される方法を用いた意味グラフを用いて、メタ関係を適用しつつ評価できる現象の例である。

【0090】

感情分析評価の例：

【0091】

20

例示的実施形態において、意味グラフは、統合医学用語システム (UMLS) オントロジ・データから構築し、又は既存のグラフを用いることができる。

【0092】

追加の医学文書を、分析する感情に関連する追加リソースとして用いることができる。医学文書には、意味グラフと同一の概念を注釈付けすることができる。固有表現語句 (named entity phrase) を、異なる文脈において固有の感情極性を有する「非常に増大した」等の感情語句と関連付けられることがある。コーパス又追加リソースにおける各文書について、ノード重み付けは、未加工値又はベクトルを用いて感情に対して構成される。

【0093】

図4は、感情分析のメタ関係重み付けを示す意味グラフからの抽出300の概略図である。意味グラフ抽出300は、UMLSの概念に関するノード310を示し、エッジ320は、ノード310間の意味関係に関する。例えば、「薬X」のノード311は、「寿命の伸長」のノード312との間に「影響を生じる」の関係321を有する。

30

【0094】

感情に対するメタ関係重み付けの数値を、ノード重み付け330及びエッジ重み付け340として示す。

【0095】

感情重み付け330、340を有する意味グラフ300を、実行時に用いて、語義曖昧性解消 (WSD) タスクに類似した、入力感情語句の正しい解釈を検出する。

【0096】

40

例示的な入力文脈は：「成分Aを薬Xに添加することにより、患者の寿命が“大きく伸長した”」対薬Yと“併用すると”薬Yの毒性が“大きく増大した”である。概念エンティティには下線が引かれ、感情は“～”で示されている。

【0097】

入力文脈を用いてグラフを活性化させる、つまり下線を引いた概念エンティティを用いて、その開始ノードにおいて意味グラフ300を活性化させ、重要な概念の各々における信号の正味蓄積は、この語句に対する感情を反映する。

【0098】

予測分析におけるバイアス / 公平性

感情分析技術の使用が普及するにつれて、感情評価における体系的なバイアスは、自然言

50

語テキストの自動感情分析に基づく決定の完全性を危うくするおそれがある。例えば、会社又その競合他社に関する記事が発行され、この記事がこれらの行為者及びその将来の財務成績について又は特定の競合製品についてのポジティブ又はネガティブな意見を表す場合、記事を適切に扱うために、いずれかの固有バイアスにさらされることは重要である。次に、記事は、決定プロセスから除外されることがあり、又はスコアを生成する際、バイアスの強度が考慮に入れられることがある。最も信頼できる記事は、いずれかの方向に偏らないものである。多くのシステムは、ノイズの多いコーパスにおける平均感情スコアは、ニュートラルなものと仮定する。

【0099】

このバイアスの検出は、複雑なプロセスであり、同一著者の以前の記事等の背景情報に依拠するが、多くの解決法は、他の文脈を無視するユーザから独立した又は局所的な手法にのみ依存する。2つのエンティティ間のバイアスの強度を有意に表す方法でのこのバイアスの集約は、記載される方法により対処される。

10

【0100】

記載される方法は、極性だけとは異なるバイアス強度の平衡表現を生成することを目標とする。所与のトピックに対する著者の公平性の測定は、データ・ソースの信頼性の評価を可能にする主要要素である。報道機関AがP社に関するポジティブな新しい記事をリリースした場合、これは中立の観点からか、又は何か既得権益があるのか？報道機関AがQ社（P社の競合他社）に関するネガティブな記事を発行した場合、これは信頼できるのか？バイアスの検出は、従来、内容の棄却をもたらす場合があるが、そのバイアスの程度及び極性を測定することができれば、そのバイアスの程度を考慮に入れることにより、内容を利用することができる。

20

【0101】

記載される方法は、バイアスを検出するだけでなく、エンティティ間の関係の極性の強度を定量化する技術も提供する。これは、NLPシステムに極めて有用である。

【0102】

組織は、通常、業務実績要約（performance summary）及び見通し（outlook）のバージョンにおいて楽観的である。予想業務実績報告が発行されると、バイアスに関して調査する必要がある。同様に、競合他社又は既得権益を有する組織が同一の会社又関連トピックに対して意見を表明すると、読み手は、同様に慎重になる。様々な通信社及び投資会社も、組織の四半期／年次報告書の評価を発行し、その多くは、組織の自己評価とは異なる。こうしたシナリオにおいて、投資家又は観察者が、現実的な又は重み付けされた平均的結論に行きつけるのは非常に有用である。これを達成するために、記載される方法は、組織と通信社等との間の関係、所有権、又は取引履歴から生じるバイアスを考慮に入れて、このバイアスを定量化することができる。

30

【0103】

自動公平調査システムにおいて、バイアス検出及び分析は、重要なタスクである。文脈（例えば、新薬、金融商品、規則等）の関連記事の選択時である。焦点の項目における信頼度の程度は、意見／意味及びその記事の文脈において検出されたバイアスの関数である。例えば、特許満了日又は政府規則の変更は、そのビジネスの立ち位置に応じて、生態系の異なるエンティティに有用又は逆効果であり得る。

40

【0104】

医学ドメインにおいて、有益な特許の満了は、保有している製薬会社にとってはビジネス上の損失であるが、この製剤のジェネリックを作りたいと思う競合他社、及びより安いコストでそれを必要とする患者にとっては利得である。同様に、政府の保護主義的関税制度は、地元生産者に利益を、輸入業者に損害をもたらし得る。報道機関又は政府機関の観点は、一方の側を他方の側に対して際立たせるか、又は中立を守るかであり、株式投資家は、イベントのバランスの取れた全体像を得て、購入／販売／保留を決定するためのツールを必要とする。

【0105】

50

予想分析におけるバイアスは、記載される方法を用いた意味グラフを用いてメタ関係を適用し、評価できる現象の例である。

【0106】

バイアス評価の例

実施形態において、意味グラフは、固有表現抽出（N E R）データを用いて、金融ドメイン・コーパスから構築することができ、以下の：

- a) 詳細な数値及びテキストによる概要、並びに短期的見通しを有する年次／四半期報告書、
- b) 金融ニュースレポート、
- c) 金融ブログ、
- d) 記事の著者、発行詳細等を含む上記全てについてのメタデータ、等のデータに関するフレームを構文解析することができる。

10

【0107】

語句レベルの感情分析をコーパス内の各文に対して行い、テキストにおいて生じるグラフ内のエンティティの各対に対して生成されるスコアを集約することができる。これは、そのドメインにおける最新の公平性のランドスケープを構成する。エンティティ間のエッジ重み付けを用いて、従来の関連付けから、極性（ポジティブ及びネガティブ）の程度を説明することができる。動的処理システムにおいては、これらのエッジ重み付けは、極性の最新の状況全体を、公平性の確率として保持することができる。

20

【0108】

バイアスは、ポジティブ、ネガティブ、又はニュートラルであり得るので、ここでの特徴ベクトルは、1又は幾つかの符号付き値（- / +）を含んで2つのエンティティ間の正味極性及び個別の極性スコアを示すことができる。

【0109】

新しい記事又文脈は、入力における同一のエンティティ及びメタデータを検出し、重み付けを用いてグラフにおける活性化拡散を実行することにより、処理することができる。信号は、エッジ重み付けの強度に従ってグラフを伝わり、極性は、特定の開始ノードを用いて、あらゆるノードに対する総和をもたらすことが、この活性化に特有である。

【0110】

次に、各ノードに対する結果としての信号蓄積値は、これらのエンティティに影響を与えるアサートが信頼できる／公平であるかどうか、又はこれらのエンティティに対する結論及びアサートが信頼できるかどうかを反映する。

30

【0111】

図5及び図6は、バイアス強度及び極性のメタ関係重み付けを示す意味グラフ400からの抽出の概略図である。

【0112】

グラフ400は、エンティティ間の意味関係についてのエッジ420を有するエンティティのノード410を含む。メタ関係重み付け430の数値は、エッジ上に示される。

【0113】

図5のノード411、412、413は、[著者A]411が著者であり、対象が[P社レポート]413及び[Q社]412である、新しい初見の文書を反映する。

40

【0114】

以下は、グラフ400の活性化をシードするのに用いられる値を反映するシナリオである。

【0115】

著者Aによる記事は、以下の内容である。：

「P社の財務見通しに関する新しいレポートは、不正確さ及び現実を全く反映しない奇妙な予測で満ちている。CEOが最新の技術に十分投資を行ったQ社とは異なり、P社は、正しい方向に進むとは思われず、このレポートは、単に、ますます悪化する業績における亀裂を取り繕う試みのように見える。」

【0116】

50

上記のこの記事の意味分析からの結果としての入力は：

著者A：重み付け + 1、ノード重み付け 4 2 1 として追加される、

P社レポート：重み付け - 1、ノード重み付け 4 2 3 として追加される、

Q社：重み付け + 1、ノード重み付け 4 2 2 として追加される。

【0117】

グラフ 4 0 0において、著者Aの前述の意見は、Aの発言の解釈に影響を与える。正味結果は、特定のトピックに関する意見は信頼すべきではないと示唆する。固有バイアス 4 5 0 の極性は、既存のエッジ重み付け 4 5 1、4 5 2、4 5 3において示される。表示バイアス 4 6 0 の極性は、エッジ重み付け 4 6 1、4 6 2、4 6 3において示される。

【0118】

特定の関係は、固有バイアスを示す。ある会社が、自分の会社について書いた記事は、会社が自らを否定的に表すものではないから、固有バイアスを有する。出版社に雇われた著者は、仕事を確保したいので、全体的立場及び出版物に対して固有バイアスを有する。会社は、競合他社に対する固有バイアスも有する。

【0119】

表示バイアスは、発行された記事の例示的シナリオなど、トピック又はエンティティに対する何らかの極性の感情のいくつかの発行されたインスタンスの反映である。

【0120】

入力（例えば、上記の記事）に続いて、信号は、重み付けされたエッジにわたって拡散する。このプロセスが一旦完了すると、活性化を分析して、信号が入力ノードにどのように蓄積したか、並びに、グラフの他のノードがどのように重み付けされるに至ったかを確認して、入力及びグラフの他のいずれかのノードに発生したノード間のバイアスを推論することができる。

【0121】

最新の記事に対するグラフ活性化は、活性化拡散値 4 7 1、4 7 2、4 7 3、4 7 4、4 7 5 を計算し、バイアス結果は以下の通りとなる。：

[P社]に対する[著者A] = (-0.9) + (-1.3) + (-1.45) = (-3.65) バイアス結果 4 7 4 として示される、

[Q社]に対する[著者A] = (+1) + (0.7) = (+1.7) バイアス結果 4 7 3 として示される。

【0122】

各値は、単一のエッジを通り、ノードに入り、開始ノードにエッジの値を掛け合わせた結果であり、エッジは、従来のバイアス（及び極性）を格納し、ノード重みは、グラフを通じて移動する信号からの正味蓄積値である。

【0123】

上記の実施形態に対する幾つかのさらに別のコメントは以下の通りである。：

著者及び出版物を表すエンティティは、従前の記事及びプレス発表等の履歴を有し、コンテンツの概念間のアサートの極性に関して分析できるコンテンツを有し得る。例として以下が挙げられる。：

「A社は破産するであろう」、又は

「B社は、AIの新機軸を開いている」、又は

「A社の財務予測は、明らかに悪化すると思われる」。

【0124】

追加の特徴は、トーン分析、感情分析、ソーシャル・ネットワーク・メタデータ（従前の雇用、顧客履歴等）、並びに共通のテーマ及びトピックのタグ付けなどの技術を用いて、こうしたコンテンツから抽出することができる。

【0125】

感情及び意見又はバイアスは、数値で定量化されることが多いが、事実に基づき、固有の極性がないように見えるが、文脈内のエンティティ間の環境関係に基づくと実際に疑問の余地がある声明に関する公平性の概念をターゲットとするのは興味深い。例えば、ある会

10

20

30

40

50

社がP社と契約し、P社の競合他社に関して、その事業に潜在的に影響を与えるかない方法で記事をリリースするとき、こうしたアサートの公平性を検知し疑問視できることが非常に重要であり、それは意味グラフの環境性質により実際に奨励される。

【0126】

グラフのエンティティと関連付けられた基準又はメタデータを含有する新しいデータがリリースされると、直接リンクされるエンティティの対の間のバイアスの程度を追跡するバックエンド・グラフ・モニタを実施することができる。どのエッジを更新すべきかの検知は、以下のような、様々な方法で行うことができる。

a) 文章、名詞句、段落、文書、又は出版物全体などの異なる範囲で同時発生するエンティティ、10

b) 例えば、親ノード 子ノード、又は同じ動詞の項のような依存関係解析構造によりリンクされるエンティティ、及び

c) テキスト内に独立した基準を有する、又は述部 (subj, obj) など、解析ツリーにおけるドメイン及び範囲エンティティを直接リンクする特定の関係によりリンクされるエンティティ。

【0127】

2つのエンティティ間のエッジ重み付けとして格納される極性についての定型句は、全ての従前のポジティブ及びネガティブ・バイアスの関数であり得るが、合計など単純なものが有効である。特定の特徴の他に対する実用性を考慮に入れる統計技術は、より頑丈でドメイン感受性重みを提供するのに有用である。20

【0128】

例えば、従来の検索 / IR、株式調査の活動等を通じた、回答コーパスに対する結果のセットを特定したら、過剰な量のバイアスを有する文書を無視すること、又はそれに応じて再重み付けすることが適切であり得る。

【0129】

1つのエンティティから他のエンティティに向けたバイアスは、ポジティブ又はネガティブ（本シナリオには、ニュートラル・バイアスは関連しない）である。著者とトピック、従前のバイアスの極性と文書コンテンツの極性の間など、2つのエンティティ間のバイアスについての分析が可能であるとき、そのデータのアライメントに従って異なる決定を行うことができる。30

【0130】

これがどのように起こるのかを示す例示的な表は以下の通りであり、同一符号のバイアス及び感情は、「関心がある (interesting)」又は「信頼できない (untrustworthy)」のいずれかとしてフラグ付けすべきであり、反対の符号のバイアス及び感情は、再重み付けすべきである。

【表1】

	バイアス	ポジティブ	ネガティブ
感情			
ポジティブ		0 *	ポジティブ**
ネガティブ		ネガティブ	0

【0131】

例えば、あるトピックに対してポジティブ・バイアスを有する著者が、何らかのポジティブな発言をする場合、それほど信頼に値しないが、何らかのネガティブな発言をする場合、バイアスのない著者の視点よりは、潜在的に注目に値すると考えられる（極性が逆の場合、逆も真である）。

【0132】

従って、上記の表において：

10

20

30

40

50

* 0 - 信頼に値せず、無視できる。

* * ポジティブ / ネガティブ - 想定されておらず、追加の値である。

ニュートラル値は、同等と考えることができる。

【 0 1 3 3 】

Alchemy Language (インターナショナル・ビジネス・マシーンズ・コーポレーションの商標) は、NLP を介したテキスト分析を与えるアプリケーション・プログラミング・インターフェースの集合である。更に別の実装においては、エンティティに関して Alchemy Language の意見の感情分析を利用することができる。

【 0 1 3 4 】

情報検索統計を用いたクエリ拡張

10

意味グラフのコンテンツに程度の差はあるが関連し得る意味分類及びエンティティに対するハードコード化重み付けの他に、意味関連性を意味グラフにおいて拡散信号として符号化する他の多くの方法がある。記載される方法の 1 つの実施形態は、ドメイン特有のコードパスから情報検索 (IR) 頻度を用いて意味関連性に対する活性化拡散の自動化構成として提供してグラフ・ノード及びリンクを重み付けすることができる。

【 0 1 3 5 】

意味関連性を考察する別 の方法は、コンテンツの特定の本体との関連の強さである。従って、関連の強さの逆数により意味グラフを再重み付けすることは、あまり一般的でない概念を増進させることになる。これは、新しいドメインに対して意味グラフを適合させるのに非常に有用である。オリジナルの意味ネットワークを構成するために用いられる初期ドメインにおいて、活性化のためのデフォルト重みは十分であり得る。しかしながら、特定の概念の関連性については、このネットワークには固有バイアスが組み込まれている。用語頻度及び同時発生頻度などの伝統的 IR 統計を用いてさえも、新しいドメインにおいてより意味のある出力を生成するために、このグラフにおいて拡散信号を適合させるのに役立ち得る。

20

【 0 1 3 6 】

追加の意味評価の例

例示的実施形態においては、例えば、UMLS - 医学ドメイン・グラフ、又は DBpedia - オープン・ドメイン・グラフなどの既存の意味グラフを用いることができる。

30

【 0 1 3 7 】

新しいドメインにおける自然言語テキストのコードパスをメタ関係の追加リソースとして用い、このドメインにおけるエンティティ及び関係のコードパス頻度を収集することができる。

【 0 1 3 8 】

値は、ノード及びエッジに対するメタ関係に関する重み付けとして集約し、意味グラフに格納することができる。

a) ノード頻度 (Fn) は、用語の一般性を反映し、従って、逆数 $1 / Fn$ が用いられる。

b) エッジ頻度 (Fe) は、用語の関連性に直接比例し、値 Fe が用いられる。

c) 0 から 1 まで又は 0 . 5 から 1 までの間に適合するようにスケールダウンされた値が、信号の伝播の再帰的減少に対する活性化拡散において非常に一般的である。

d) 上記のその他の変形は、ノード及びエッジに対して tf / idf スコアを使用することを含む。

40

【 0 1 3 9 】

入力文脈が与えられると、エンティティのインスタンスが意味グラフにおいて認識され、グラフはその点において活性化され、活性化拡散が、語義曖昧性解消 (WSD) 、語義推定 (WSI) 又はクエリ拡張などの特定の NLP タスクの文脈において実行され得る。

【 0 1 4 0 】

出力は、その他の場合においてより、はるかにドメインに関連する。

【 0 1 4 1 】

クエリ拡張の例

クエリ拡張のプロセスは、入力語句又はクエリの文脈を理解すること、及び、オリジナル

50

のクエリに対する適正文書を見つけるのに役立つ代替的な又は追加の語句を見つけることであると考えることができる。1つの例は、「A社の事務所はアイルランドのどこにあるか?」である。「アイルランド (Ireland)」を、A社が事務所を有する「コーク (Cork)」、「ダブリン (Dublin)」及び「ガルウェイ (Galway)」に改めると、正しい回答を見つける可能性が潜在的に向上する。通常、クエリ拡張は、入力の観点に焦点を当て、合致の尤度を向上させるように補外を試みる。

【0142】

本技術をクエリ拡張において適用する効果を実証することができる。中央ノードは、クエリ拡張の候補であり、周囲のノードは、入力文脈からのものであり得る。情報検索統計を考慮に入れるように活性化拡散を構成した後、より一般的概念の関連度スコアは、十分に低くなり、文脈に関連のないものとしてそれらを除外できる。この波及効果は、クエリ拡張の候補のランクも変化することである。ランクが相互に関連する関連入力の数によるものであるとき、新しい上位ランクのクエリ拡張候補を得ることができる。

10

【0143】

入力「公記録」、「法律」、「情報」、「行政的小区域」等に対して追加の有用なクエリ用語を得ようとする初期の試みにより、「法執行機関」のような中央ノードの最良の候補を得ることができる。ドメイン・コーパスにおけるノードの頻度に従ってグラフの重み付けを行った後、オリジナルの入力の幾つかは、もはや適切な入力とはみなされなくなることがあり、代わりに「秘密情報」という用語が中央ノードとして提案され得る。オリジナルの文脈が「データ・プライバシー」というドメインにあったとすると、これは大きな改善である。

20

【0144】

自動リスク評価

自動リスク評価の分野は、所与のキーワードと特定のレベルのリスク危険度との間に直接リンクがあるキーワード・ベースの手法に主として依存する。機械学習技術は、通常、キーワードとそれに関連するリスク要因に基づくより頑強な認識を得るのに用いる。リスク危険度を推定し、文脈に直ちには存在しないかもしれないという、知見を用いて定量化するアプローチは、既存の技術に対して有意な追加価値となる。

【0145】

リスクに対する危険度の程度つまり強度は、感情分析又は地理空間推定と類似した方法で、特定のリスク要因に関して定量化することができるものである。同一技術を利用してリスク評価を実行する方法の更に別の実施形態が期待される。NLP技術を用いるこの問題への既存の解決法は、語彙ベース感情分析と同様に、特定のキーワード又は語句をリスクに関連付けるものである。より環境的な方法でリスク危険度を推定できるシステムは、文脈におけるリスク危険度に対する強度尺度を生成するのに極めて有用である。

30

【0146】

医学ドメインから例を一つ挙げると、患者の心臓血管系疾患を発症する10年リスクを推測するためのFraminghamリスク評価の詳細を、グラフ構造に符号化することができる。グラフにおいて活性化拡散を実行すると、関連する固有エンティティの各々への信号伝播の強度は、患者のリスクのレベルに直接比例するようになる。

40

【0147】

地理空間推定

バイアスを歴史的感情極性の強度に従って定量化できるのと同じ方法で、エンティティ又はトピックの地理的場所のリストとの関連性を、歴史的関連付け及び地理的距離に従って評価することができる。地理空間推定の既存の方法は、テキストにおける直接的言及に限定されるが、テキスト内に複数の位置が存在する場合、ドメイン・コーパスにおける普遍的関連付けにより関連性を決定できる方法が、はるかに有用である。

【0148】

地理的関連性に対するメタ関係重み付けの特徴ベクトルは、以下の特徴のサンプル特徴の組み合わせ、すなわち：

50

- a) 2 つのエンティティ間の正規化した地理的距離、
- b) 意味的関連付けの関連性（例えば、オリジナルの活性化拡散実装による重み）、
- c) 地理空間意味分類のみが伝達され、重みの蓄積に使用される、b) の代案、
を含むことができる。

【 0 1 4 9 】

例えば、カリフォルニア州立法府に固有の多くの語彙を用いるが、他の州の法を引用する又はカリフォルニア州にはない特定の場所の名を挙げるに、カリフォルニア州法令の中の法を、問題の法のために（合衆国（ U S A ）と欧州連合（ E U ）ベースの企業間での顧客データ伝送など）取り上げる。カリフォルニア州とテキスト内に明示的に言及される場所との関連性を推定できる環境技術は、法的ドメイン N L P タスクに極めて有用である。

10

【 0 1 5 0 】

ここで図 7 を参照すると、コンピュータ・システム又はサーバの形態のシステム 5 0 0 の例の概略が示される。

【 0 1 5 1 】

コンピュータ・システム又はサーバ 5 1 2 は、多数の他の汎用又は専用コンピューティング・システム環境又は構成と共に動作することができる。コンピュータ・システム / サーバ 5 1 2 と共に使用するのに好適であり得る周知のコンピューティング・システム、環境、及び / 又は構成の例として、これらに限定されるものではないが、パーソナル・コンピュータ・システム、サーバ・コンピュータ・システム、シン・クライアント、シック・クライアント、手持ち式又はラップトップ型デバイス、マルチプロセッサ・システム、マイクロプロセッサ・ベースのシステム、セットトップボックス、プログラム可能民生電子機器、ネットワーク P C 、ミニコンピュータ・システム、メインフレーム・コンピュータ・システム、及び、上述のシステム又はデバイスのいずれかを含む分散型クラウド・コンピューティング環境等が含まれる。

20

【 0 1 5 2 】

コンピュータ・システム / サーバ 5 1 2 は、コンピュータ・システムによって実行される、プログラム・モジュールなどのコンピュータ・システム実行可能命令の一般的な文脈で説明することができる。一般に、プログラム・モジュールは、特定のタスクを実行する又は特定の抽象データ型を実装する、ルーチン、プログラム、オブジェクト、コンポーネント、論理、データ構造などを含むことができる。コンピュータ・システム / サーバ 5 1 2 は、通信ネットワークを通じてリンクされた遠隔処理デバイスによってタスクが実行される分散型クラウド・コンピューティング環境で実施することができる。分散型クラウド・コンピューティング環境において、プログラム・モジュールは、メモリ・ストレージ・デバイスを含む、ローカル及び遠隔両方のコンピュータ・システム・ストレージ媒体内に配置することができる。

30

【 0 1 5 3 】

図 7 において、コンピュータ・システム / サーバ 5 1 2 は、汎用コンピューティング・デバイスの形で示される。コンピュータ・システム / サーバ 5 1 2 のコンポーネントは、これらに限定されるものではないが、1 つ又は複数のプロセッサ又は処理ユニット 5 1 6 、システム・メモリ 5 2 8 、及びシステム・メモリ 5 2 8 を含む種々のシステム・コンポーネントをプロセッサ 5 1 6 に結合するバス 5 1 8 を含むことができる。

40

【 0 1 5 4 】

バス 5 1 8 は、メモリ・バス又はメモリ・コントローラ、周辺バス、アクセラレーテッド・グラフィックス・ポート、及び種々のバス・アーキテクチャのいずれかを用いるプロセッサ又はローカル・バスを含む、幾つかのタイプのバス構造のうちのいずれかの1 つ又は複数を表す。限定ではなく例としては、このようなアーキテクチャは、業界標準アーキテクチャ（ Industry Standard Architecture 、 I S A ）バス、マイクロ・チャネル・アーキテクチャ（ Micro Channel Architecture 、 M C A ）バス、 Enhanced I S A （ E I S A ）バス、 Video E l e c t r o n i c s S t a n d a r d s A s s o c i a t i o n （ V E S A ）ローカル・バス、及び Peripheral Compon

50

ent Interconnects (PCI) バスを含む。

【0155】

コンピュータ・システム / サーバ 512 は、典型的には、種々のコンピュータ・システム可読媒体を含む。このような媒体は、コンピュータ・システム / サーバ 512 によりアクセス可能ないずれかの利用可能媒体とすることことができ、揮発性媒体及び不揮発性媒体の両方と、取り外し可能媒体及び取り外し不能媒体の両方とを含む。

【0156】

システム・メモリ 528 は、ランダム・アクセス・メモリ (RAM) 530 及び / 又はキヤッシュ・メモリ 532 など、揮発性メモリの形のコンピュータ・システム可読媒体を含むことができる。コンピュータ・システム / サーバ 512 は、他の取り外し可能 / 取り外し不能、揮発性 / 不揮発性のコンピュータ・システム・ストレージ媒体をさらに含むことができる。単なる例として、取り外し不能の不揮発性磁気媒体 (図示されておらず、典型的には「ハード・ドライブ」と呼ばれる)との間の読み出し及び書き込みのために、ストレージ・システム 534 を設けることができる。図示されていないが、取り外し可能な不揮発性磁気ディスク (例えば、「フロッピー・ディスク」)との間の読み出し及び書き込みのための磁気ディスク・ドライブと、CD - ROM、DVD - ROM 又は他の光媒体などの取り外し可能な不揮発性光ディスクとの間の読み出し及び書き込みのための光ディスク・ドライブとを設けることができる。このような例においては、それぞれを、1つ又は複数のデータ媒体インターフェースによってバス 518 に接続することができる。以下でさらに示され説明されるように、メモリ 528 は、本発明の実施形態の機能を実行するように構成されたプログラム・モジュールのセット (例えば、少なくとも 1 つ) を有する少なくとも 1 つのプログラム製品を含むことができる。

10

20

30

【0157】

限定ではなく例として、メモリ 528 内に、プログラム・モジュール 542 のセット (少なくとも 1 つ) を有するプログラム / ユーティリティ 540、並びにオペレーティング・システム、1つ又は複数のアプリケーション・プログラム、他のプログラム・モジュール、及びプログラム・データを格納することができる。オペレーティング・システム、1つ又は複数のアプリケーション・プログラム、他のプログラム・モジュール、及びプログラム・データ、又はそれらの何らかの組み合わせの各々は、ネットワーキング環境の実装を含むことができる。プログラム・モジュール 542 は、一般に、本明細書で説明される本発明の実施形態の機能及び / 又は方法を実行する。

30

【0158】

コンピュータ・システム / サーバ 512 は、キーボード、ポインティング・デバイス、ディスプレイ 524 等のような 1 つ又は複数の外部デバイス 514 ; ユーザがコンピュータ・システム / サーバ 512 と対話することを可能にする 1 つ又は複数のデバイス ; 及び / 又はコンピュータ・システム / サーバ 512 が 1 つ又は複数の他のコンピューティング・デバイスと通信することを可能にするいずれかのデバイス (例えば、ネットワーク・カード、モデムなど) と通信することもできる。このような通信は、入力 / 出力 (I/O) インターフェース 522 を経由して行うことができる。さらにまた、コンピュータ・システム / サーバ 512 は、ネットワーク・アダプタ 520 を介して、ローカル・エリア・ネットワーク (LAN)、汎用広域ネットワーク (WAN)、及び / 又はパブリック・ネットワーク (例えば、インターネット) などの 1 つ又は複数のネットワークと通信することもできる。示されるように、ネットワーク・アダプタ 520 は、バス 518 を介して、コンピュータ・システム / サーバ 512 の他のコンポーネントと通信する。図示されていないが、コンピュータ・システム / サーバ 512 と共に他のハードウェア及び / 又はソフトウェア・コンポーネントを使用できることを理解されたい。例としては、これらに限定されるものではないが、マイクロコード、デバイス・ドライバ、冗長処理ユニット、外部ディスク・ドライブ・アレイ、RAID システム、テープ・ドライブ、及びデータ・アーカイブ・ストレージ・システムなどが含まれる。

40

【0159】

50

本発明は、統合のいずれかの可能な技術的詳細レベルにおける、システム、方法、及び／又はコンピュータ・プログラム製品とすることができる。コンピュータ・プログラム製品は、プロセッサに本発明の態様を実行させるためのコンピュータ可読プログラム命令をその上に有するコンピュータ可読ストレージ媒体（単数又は複数）を含むことができる。

【 0 1 6 0 】

コンピュータ可読ストレージ媒体は、命令実行デバイスにより使用される命令を保持及び格納できる有形デバイスとすることができます。コンピュータ可読ストレージ媒体は、例えば、これらに限定されるものではないが、電子ストレージ・デバイス、磁気ストレージ・デバイス、光ストレージ・デバイス、電磁気ストレージ・デバイス、半導体ストレージ・デバイス、又は上記のいずれかの適切な組み合わせとすることができます。コンピュータ可読ストレージ媒体のより具体的な例の非網羅的なリストとして、以下のもの：すなわち、ポータブル・コンピュータ・ディスクケット、ハードディスク、ランダム・アクセス・メモリ（R A M）、読み出し専用メモリ（R O M）、消去可能プログラム可能読み出し専用メモリ（E P R O M又はフラッシュ・メモリ）、スタティック・ランダム・アクセス・メモリ（S R A M）、ポータブル・コンパクト・ディスク読み出し専用メモリ（C D - R O M）、デジタル多用途ディスク（D V D）、メモリ・スティック、フロッピー・ディスク、パンチカード若しくは命令がそこに記録された溝内の隆起構造のような機械的にエンコードされたデバイス、及び上記のいずれかの適切な組み合わせが挙げられる。本明細書で使用される場合、コンピュータ可読ストレージ媒体は、電波、又は他の自由に伝搬する電磁波、導波管若しくは他の伝送媒体を通じて伝搬する電磁波（例えば、光ファイバ・ケーブルを通る光パルス）、又はワイヤを通って送られる電気信号などの、一時的信号自体として解釈されない。

10

20

30

【 0 1 6 1 】

本明細書で説明されるコンピュータ可読プログラム命令は、コンピュータ可読ストレージ媒体からそれぞれのコンピューティング／処理デバイスにダウンロードすることができ、又は、例えばインターネット、ローカル・エリア・ネットワーク、広域ネットワーク、及び／又は無線ネットワークなどのネットワークを介して外部コンピュータ又は外部ストレージ・デバイスにダウンロードすることができる。ネットワークは、銅伝送ケーブル、光伝送ファイバ、無線伝送、ルータ、ファイアウォール、スイッチ、ゲートウェイ・コンピュータ、及び／又はエッジ・サーバを含むことができる。各コンピューティング／処理デバイスにおけるネットワーク・アダプタ・カード又はネットワーク・インターフェースは、ネットワークからコンピュータ可読プログラム命令を受け取り、それぞれのコンピューティング／処理デバイス内のコンピュータ可読ストレージ媒体内に格納するためにコンピュータ可読プログラム命令を転送する。

【 0 1 6 2 】

本発明の動作を実行するためのコンピュータ可読プログラム命令は、アセンブラー命令、命令セットアーキテクチャ（I S A）命令、マシン命令、マシン依存命令、マイクロコード、ファームウェア命令、状態設定データ、又は、S m a l l t a l k、C + +などのオブジェクト指向プログラミング言語、又は、「C」プログラミング言語若しくは類似のプログラミング言語などの通常の手続き型プログラミング言語を含む1つ又は複数のプログラミング言語の任意の組み合わせで記述することができるいずれかのソースコード若しくはオブジェクトコードとすることができます。コンピュータ可読プログラム命令は、完全にユーザのコンピュータ上で実行される場合もあり、一部がユーザのコンピュータ上で、独立型ソフトウェア・パッケージとして実行される場合もあり、一部がユーザのコンピュータ上で実行され、一部が遠隔コンピュータ上で実行される場合もあり、又は完全に遠隔コンピュータ若しくはサーバ上で実行される場合もある。最後のシナリオにおいて、遠隔コンピュータは、ローカル・エリア・ネットワーク（L A N）若しくは広域ネットワーク（W A N）を含むいずれかのタイプのネットワークを通じてユーザのコンピュータに接続される場合もあり、又は外部コンピュータへの接続がなされる場合もある（例えば、インターネットサービスプロバイダを用いたインターネットを通じて）。幾つかの実施形態におい

40

50

て、例えば、プログラム可能論理回路、フィールド・プログラマブル・ゲート・アレイ（FPGA）、又はプログラム可能論理アレイ（PLA）を含む電子回路は、コンピュータ可読プログラム命令の状態情報を用いて電子回路を個人化することによりコンピュータ可読プログラム命令を実行し、本発明の態様を実施することができる。

【0163】

本発明の態様は、本発明の実施形態による方法、装置（システム）及びコンピュータ・プログラム製品のフローチャート図及び／又はブロック図を参照して説明される。フローチャート図及び／又はブロック図の各ブロック、並びにフローチャート図及び／又はブロック図内のブロックの組み合わせは、コンピュータ可読プログラム命令によって実装できることが理解されるであろう。

10

【0164】

これらのコンピュータ可読プログラム命令を、汎用コンピュータ、専用コンピュータ、又は他のプログラム可能データ処理装置のプロセッサに与えてマシンを製造し、それにより、コンピュータ又は他のプログラム可能データ処理装置のプロセッサによって実行される命令が、フローチャート及び／又はブロック図の1つ又は複数のブロック内で指定された機能／動作を実装するための手段を作り出すようにすることができる。これらのコンピュータ可読プログラム命令を、コンピュータ、プログラム可能データ処理装置、及び／又は他のデバイスを特定の方式で機能させるように指示することができるコンピュータ可読媒体内に格納し、それにより、命令が内部に格納されたコンピュータ可読ストレージ媒体が、フローチャート及び／又はブロック図の1つ又は複数のブロックにおいて指定された機能／動作の態様を実装する命令を含む製品を含むようにすることもできる。

20

【0165】

コンピュータ可読プログラム命令を、コンピュータ、他のプログラム可能データ処理装置、又は他のデバイス上にロードして、一連の動作ステップをコンピュータ、他のプログラム可能データ処理装置、又は他のデバイス上で行わせてコンピュータ実装プロセスを生成し、それにより、コンピュータ又は他のプログラム可能装置上で実行される命令が、フローチャート及び／又はブロック図の1つ又は複数のブロックにおいて指定された機能／動作を実装するようにすることもできる。

【0166】

図面内のフローチャート及びブロック図は、本発明の種々の実施形態による、システム、方法、及びコンピュータ・プログラム製品の可能な実装の、アーキテクチャ、機能及び動作を示す。この点に関して、フローチャート内の各ブロックは、指定された論理機能を実装するための1つ又は複数の実行可能命令を含む、モジュール、セグメント、又はコードの一部を表すことができる。幾つかの代替的な実装において、ブロック内に示される機能は、図に示される順序とは異なる順序で生じることがある。例えば、連続して示される2つのブロックは、関与する機能に応じて、実際には実質的に同時に実行されることもあり、又はこれらのブロックはときとして逆順で実行されることもある。ブロック図及び／又はフローチャート図の各ブロック、及びブロック図及び／又はフローチャート図内のブロックの組み合わせは、指定された機能又は動作を実行する、又は専用のハードウェアとコンピュータ命令との組み合わせを実行する、専用ハードウェア・ベースのシステムによって実装できることにも留意されたい。

30

【0167】

クラウド・コンピューティング

本開示はクラウド・コンピューティングについての詳細な説明を含むが、本明細書に記載される教示の実装は、クラウド・コンピューティング環境に限定されないことを理解されたい。むしろ、本発明の実施形態は、現在既知の又は後で開発される他のいずれかのタイプのコンピューティング環境と共に実施することができる。

【0168】

クラウド・コンピューティングは、最小限の管理労力又はサービス・プロバイダとの対話で迅速にプロビジョニング及び解放することができる構成可能なコンピューティング・リ

40

50

ソース（例えば、ネットワーク、ネットワーク帯域幅、サーバ、処理、メモリ、ストレージ、アプリケーション、仮想マシン、及びサービス）の共有プールへの、便利なオンデマンドのネットワーク・アクセスを可能にするためのサービス配信のモデルである。このクラウド・モデルは、少なくとも5つの特徴、少なくとも3つのサービス・モデル、及び少なくとも4つの配備モデルを含むことができる。

【0169】

特徴は、以下の通りである。

オンデマンド・セルフサービス：クラウド・コンシューマは、必要に応じて、サーバ時間及びネットワーク・ストレージ等のコンピューティング機能を、人間がサービスのプロバイダと対話する必要なく自動的に、一方的にプロビジョニングすることができる。10

広範なネットワーク・アクセス：機能は、ネットワーク上で利用可能であり、異種のシン又はシック・クライアント・プラットフォーム（例えば、携帯電話、ラップトップ、及びPDA）による使用を促進する標準的な機構を通じてアクセスされる。

リソースのプール化：プロバイダのコンピューティング・リソースは、マルチ・テナント・モデルを用いて、異なる物理及び仮想リソースを要求に応じて動的に割り当て及び再割り当てすることにより、複数のコンシューマにサービスを提供するためにプールされる。コンシューマは、一般に、提供されるリソースの正確な位置についての制御又は知識を持たないが、より高レベルの抽象化では位置（例えば、国、州、又はデータセンタ）を特定できる場合があるという点で、位置とは独立しているといえる。

迅速な弾力性：機能は、迅速かつ弾力的に、場合によっては自動的に、プロビジョニングして素早くスケール・アウトし、迅速にリリースして素早くスケール・インさせることができる。コンシューマにとって、プロビジョニングに利用可能なこれらの機能は、多くの場合、無制限であり、いつでもどんな量でも購入できるように見える。20

計測されるサービス：クラウド・システムは、サービスのタイプ（例えば、ストレージ、処理、帯域幅、及びアクティブなユーザ・アカウント）に適した何らかの抽象化レベルでの計量機能を用いることによって、リソースの使用を自動的に制御及び最適化する。リソース使用を監視し、制御し、報告し、利用されるサービスのプロバイダとコンシューマの両方に対して透明性をもたらすことができる。

【0170】

サービス・モデルは以下の通りである。

S o f t w a r e a s a S e r v i c e (S a a S)：クラウド・インフラストラクチャ上で動作しているプロバイダのアプリケーションを使用するために、コンシューマに提供される機能である。これらのアプリケーションは、ウェブ・ブラウザ（例えば、ウェブ・ベースの電子メール）などのシン・クライアント・インターフェースを通じて、種々のクライアント・デバイスからアクセス可能である。コンシューマは、限定されたユーザ固有のアプリケーション構成設定の考え得る例外として、ネットワーク、サーバ、オペレーティング・システム、ストレージ、又は個々のアプリケーション機能をも含めて、基礎をなすクラウド・インフラストラクチャを管理又は制御しない。30

P l a t f o r m a s a S e r v i c e (P a a S)：プロバイダによってサポートされるプログラミング言語及びツールを用いて生成された、コンシューマが生成した又は取得したアプリケーションを、クラウド・インフラストラクチャ上に配備するために、コンシューマに提供される機能である。コンシューマは、ネットワーク、サーバ、オペレーティング・システム、又はストレージなどの基礎をなすクラウド・インフラストラクチャを管理又は制御しないが、配備されたアプリケーション、及び場合によってはアプリケーション・ホスティング環境構成に対して制御を有する。40

I n f r a s t r u c t u r e a s a S e r v i c e (I a a S)：コンシューマが、オペレーティング・システム及びアプリケーションを含み得る任意のソフトウェアを配備及び動作させることができる、処理、ストレージ、ネットワーク、及び他の基本的なコンピューティング・リソースをプロビジョニングするため、コンシューマに提供される機能である。コンシューマは、基礎をなすクラウド・インフラストラクチャを管理又は50

制御しないが、オペレーティング・システム、ストレージ、配備されたアプリケーションに対する制御、及び場合によってはネットワーク・コンポーネント（例えば、ホストのファイアウォール）選択の限定された制御を有する。

【0171】

配備モデルは以下の通りである。

プライベート・クラウド：クラウド・インフラストラクチャは、ある組織のためだけに運営される。このクラウド・インフラストラクチャは、その組織又は第三者によって管理することができ、オンプレミス又はオフプレミスに存在することができる。

コミュニティ・クラウド：クラウド・インフラストラクチャは、幾つかの組織によって共有され、共通の関心事項（例えば、任務、セキュリティ要件、ポリシー、及びコンプライアンス上の考慮事項）を有する特定のコミュニティをサポートする。クラウド・インフラストラクチャは、その組織又は第三者によって管理することができ、オンプレミス又はオフプレミスに存在することができる。

10

パブリック・クラウド：クラウド・インフラストラクチャは、一般公衆又は大規模な業界グループに利用可能であり、クラウド・サービスを販売する組織によって所有される。

ハイブリッド・クラウド：クラウド・インフラストラクチャは、固有のエンティティのままであるが、データ及びアプリケーションの移行性を可能にする標準化された又は専用の技術（例えば、クラウド間の負荷分散のためのクラウド・バースティング）によって結び付けられる2つ以上のクラウド（プライベート、コミュニティ、又はパブリック）の混成物である。

20

【0172】

クラウド・コンピューティング環境は、無国籍性、低結合性、モジュール性、及びセマンティック相互運用性に焦点を置くことを指向するサービスである。クラウド・コンピューティングの中心は、相互接続されたノードのネットワークを含むインフラストラクチャである。

【0173】

ここで図8を参照すると、例示的なクラウド・コンピューティング環境650が示される。示されるように、クラウド・コンピューティング環境650は、例えばパーソナル・デジタル・アシスタント（PDA）又は携帯電話654A、デスクトップ・コンピュータ654B、ラップトップ・コンピュータ654C、及び/又は自動車コンピュータ・システム654Nなどといった、クラウド・コンシューマによって用いられるローカル・コンピューティング・デバイスと通信することができる、1つ又は複数のクラウド・コンピューティング・ノード610を含む。ノード610は、互いに通信することができる。これらのノードは、上述のようなプライベート・クラウド、コミュニティ・クラウド、パブリック・クラウド、若しくはハイブリッド・クラウド、又はこれらの組み合わせなど、1つ又は複数のネットワークにおいて物理的又は仮想的にグループ化することができる（図示せず）。これにより、クラウド・コンピューティング環境650は、クラウド・コンシューマがローカル・コンピューティング・デバイス上にリソースを保持する必要のないサービスとして、インフラストラクチャ、プラットフォーム、及び/又はソフトウェアを提供することができる。図8に示されるコンピューティング・デバイス654A～Nのタイプは単に例示であることを意図し、コンピューティング・ノード610及びクラウド・コンピューティング環境650は、いずれのタイプのネットワーク及び/又はネットワーク・アドレス指定可能な接続上でも（例えば、ウェブ・プラウザを用いて）、いずれのタイプのコンピュータ化された装置とも通信できることを理解されたい。

30

【0174】

ここで図9を参照すると、クラウド・コンピューティング環境650（図8）によって提供される機能抽象化層の組が示される。図9に示されるコンポーネント、層、及び機能は単に例示であることを意図し、本発明の実施形態はそれらに限定されないことを予め理解されたい。図示されるように、以下の層及び対応する機能が提供される。：

【0175】

40

50

ハードウェア及びソフトウェア層 7 6 0 は、ハードウェア及びソフトウェア・コンポーネントを含む。ハードウェア・コンポーネントの例として、メインフレーム 7 6 1、R I S C (Reduced Instruction Set Computer (縮小命令セット・コンピュータ)) アーキテクチャ・ベースのサーバ 7 6 2、サーバ 7 6 3、ブレード・サーバ 7 6 4、ストレージ・デバイス 7 6 5、並びにネットワーク及びネットワーキング・コンポーネント 7 6 6 が含まれる。幾つかの実施形態において、ソフトウェア・コンポーネントは、ネットワーク・アプリケーション・サーバ・ソフトウェア 7 6 7 及びデータベース・ソフトウェア 7 6 8 を含む。

【 0 1 7 6 】

仮想化層 7 7 0 は、抽象化層を提供し、この層により、仮想エンティティの以下の例、すなわち、仮想サーバ 7 7 1、仮想ストレージ 7 7 2、仮想プライベート・ネットワークを含む仮想ネットワーク 7 7 3、仮想アプリケーション及びオペレーティング・システム 7 7 4、並びに仮想クライアント 7 7 5 を提供することができる。

10

【 0 1 7 7 】

一例においては、管理層 7 8 0 は、以下で説明される機能を提供することができる。リソース・プロビジョニング 7 8 1 は、クラウド・コンピューティング環境内でタスクを実行するために利用されるコンピューティング・リソース及び他のリソースの動的な調達を提供する。計量及び価格決定 7 8 2 は、クラウド・コンピューティング環境内でリソースが利用される際のコスト追跡と、これらのリソースの消費に対する課金又は請求とを提供する。一例においては、これらのリソースは、アプリケーション・ソフトウェア・ライセンスを含むことができる。セキュリティは、クラウド・コンシューマ及びタスクに対する識別情報の検証と、データ及び他のリソースに対する保護とを提供する。ユーザ・ポータル 8 8 3 は、コンシューマ及びシステム管理者のために、クラウド・コンピューティング環境へのアクセスを提供する。サービス・レベル管理 8 8 4 は、要求されるサービス・レベルが満たされるように、クラウド・コンピューティング・リソースの割り当て及び管理を提供する。サービス・レベル・アグリーメント (Service Level Agreement、S L A) の計画及び履行 7 8 5 は、S L A に従って将来の要件が予測されるクラウド・コンピューティング・リソースの事前配置及び調達を提供する。

20

【 0 1 7 8 】

ワークロード層 7 9 0 は、クラウド・コンピューティング環境を利用することができる機能の例を提供する。この層から提供することができるワークロード及び機能の例には、マッピング及びナビゲーション 7 9 1、ソフトウェア開発及びライフサイクル管理 7 9 2、仮想教室教育配信 7 9 3、データ分析処理 7 9 4、トランザクション処理 7 9 5、及びメッセージ変換処理 7 9 6 が含まれる。

30

【 0 1 7 9 】

本発明の種々の実施形態の説明は、例証の目的のために提示されたが、これらは、網羅的であること、又は本発明を開示した実施形態に限定することを意図するものではない。当業者には、説明される実施形態の範囲及び趣旨から逸脱することなく、多くの修正及び変形が明らかであろう。本明細書で用いられる用語は、実施形態の原理、実際の適用、又は市場に見られる技術に優る技術的改善を最もよく説明するため、又は、当業者が、本明細書に開示される実施形態を理解するのを可能にするために選択された。

40

【 0 1 8 0 】

本発明の範囲から逸脱することなく、上記に対する改善及び修正を行うことができる。

【 符号の説明 】

【 0 1 8 1 】

1 0 0、1 5 0 : 流れ図

1 1 0 : 意味リソース

1 1 1、1 1 5 : 意味グラフ

1 1 2 : 追加リソース

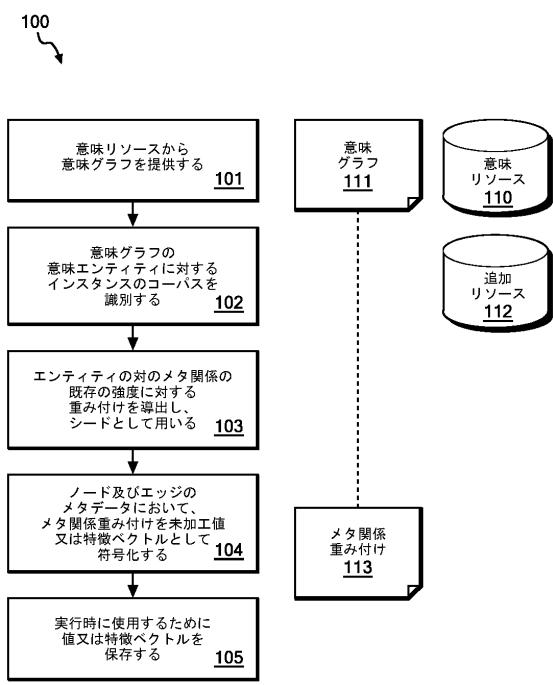
1 1 3、1 1 6 : メタ関係重み付け

50

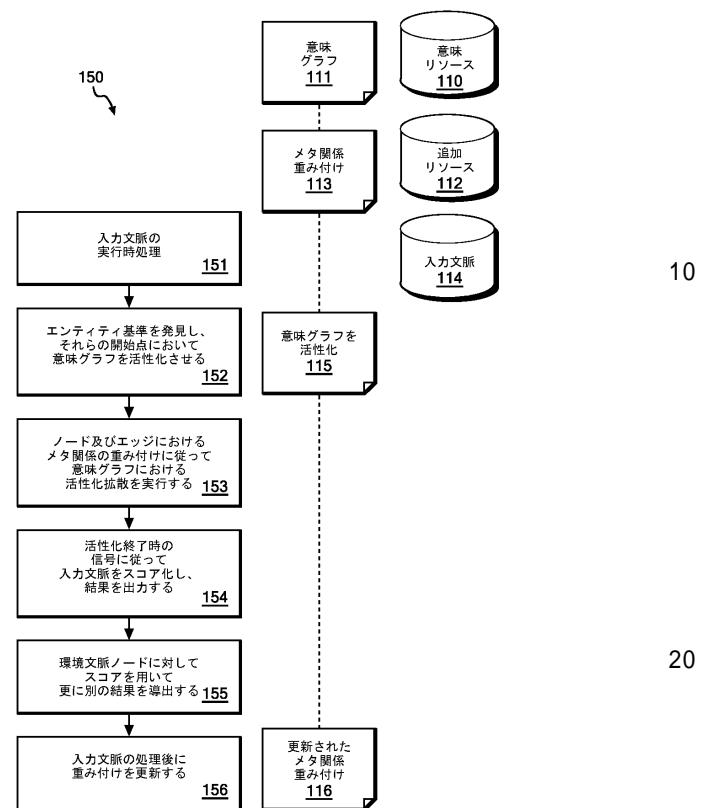
1 1 4 : 入力文脈	
2 0 0 : システム	
2 0 1 : プロセッサ	
2 0 2 : メモリ	
2 0 3 : コンピュータ命令	
2 1 0 : 意味グラフ評価システム	
2 1 0、5 1 6 : プロセッサ	
2 2 0 : メタ関係コンポーネント	
2 2 2 : シード重み付け導出コンポーネント	10
2 2 3 : 重み付け追加コンポーネント	
2 2 4 : 重み付け更新コンポーネント	
2 3 0 : 実行時コンポーネント	
2 3 1 : 入力文脈選択コンポーネント	
2 3 2 : 概念検出コンポーネント	
2 3 3 : グラフ活性化コンポーネント	
2 3 4 : スコア化コンポーネント	
2 3 5 : 環境文脈コンポーネント	
2 3 7 : 出力コンポーネント	
3 0 0、4 0 0 : 意味グラフ	
3 1 0、3 1 2、4 1 0、4 1 1、4 1 2、4 1 3 : ノード	20
3 2 0、4 2 0 : エッジ	
4 2 1、4 2 2、4 2 3 : ノード重み付け	
4 5 0 : 固有バイアス	
4 5 1、4 5 2、4 5 3、4 6 1、4 6 2、4 6 3 : エッジ重み付け	
4 6 0 : 表示バイアス	
4 7 1、4 7 2、4 7 3、4 7 4、4 7 5 : 活性化拡散値	
5 1 2 : コンピュータ・システム／サーバ	
5 2 8 : システム・メモリ	
6 5 0 : クラウド・コンピューティング環境	
7 6 0 : ハードウェア及びソフトウェア層	30
7 6 0 : 仮想化層	
7 8 0 : 管理層	
7 9 0 : ワークロード層	

【図面】

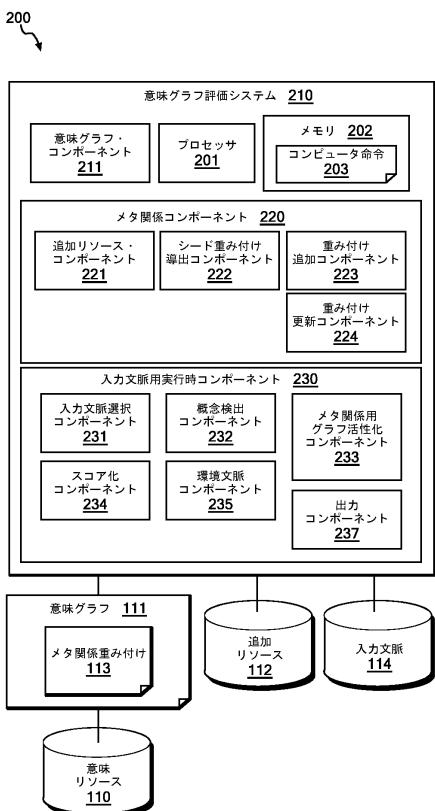
【図 1】



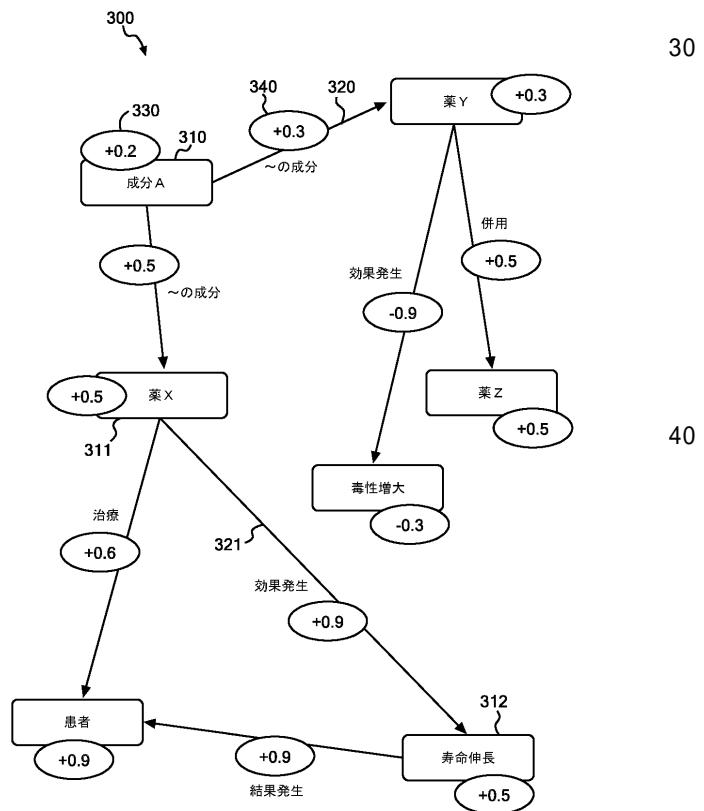
【図 2】



【図 3】



【図 4】



10

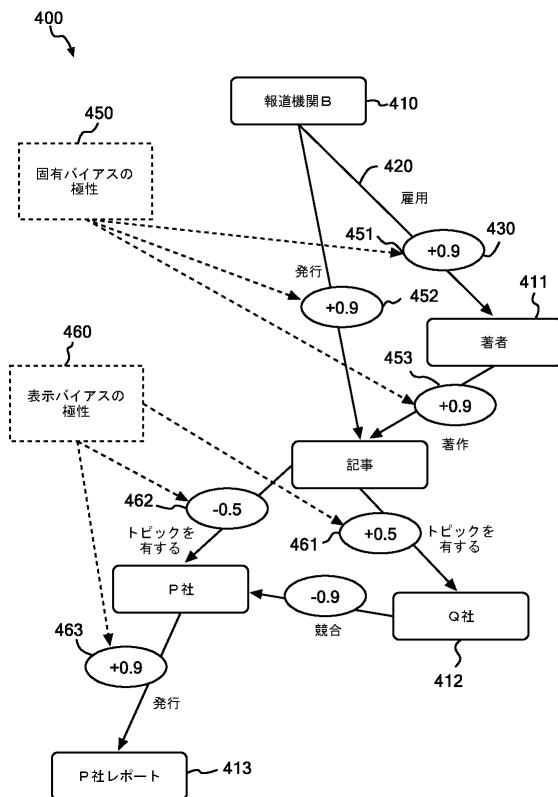
20

30

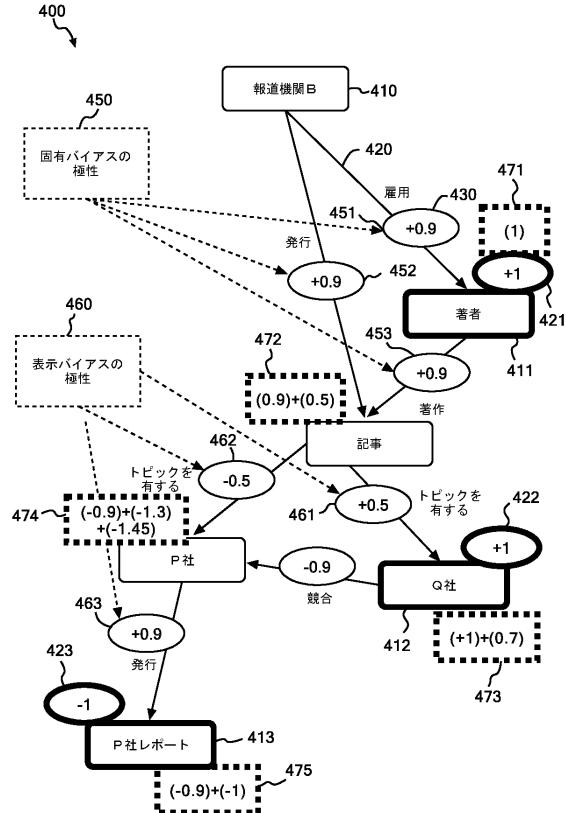
40

50

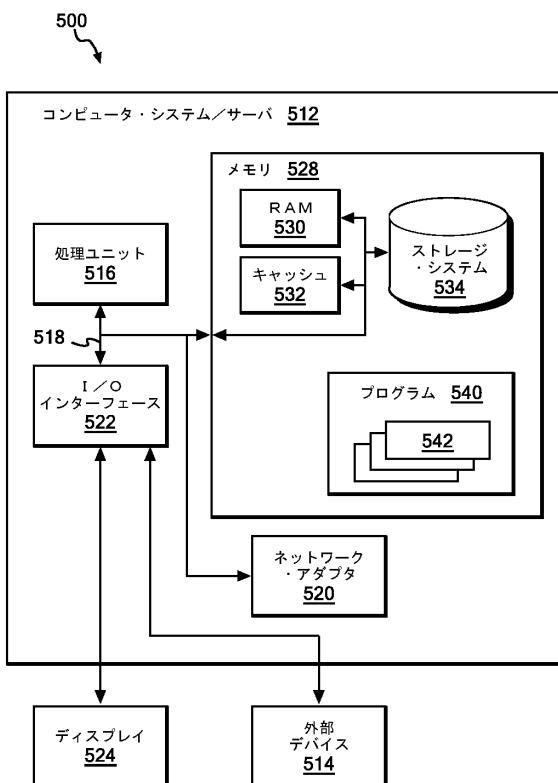
【図5】



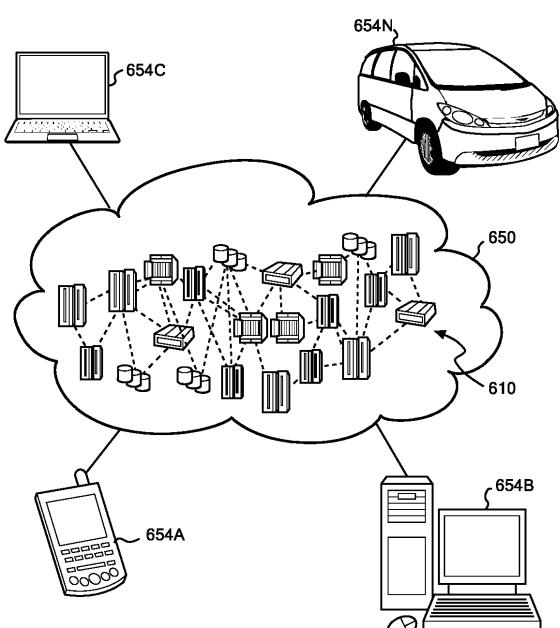
【 四 6 】



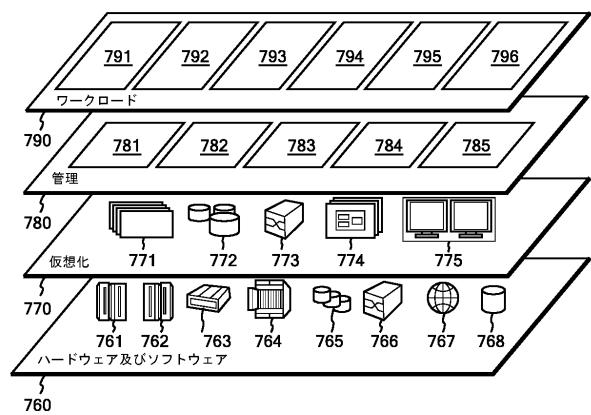
【図7】



【 四 8 】



【図9】



10

20

30

40

50

フロントページの続き

(33) 優先権主張国・地域又は機関

米国(US)

(74) 代理人 100112690

弁理士 太佐 種一

(72) 発明者 モハン、アディティア

アイルランド共和国 15 モルハダート ダマスタウン・インダストリアル・エstate テクノロジー・キャンパス

(72) 発明者 マカティア、シーマス

アイルランド共和国 15 モルハダート ダマスタウン・インダストリアル・エstate テクノロジー・キャンパス

(72) 発明者 マクロスキー、ダニエル

アイルランド共和国 15 モルハダート ダマスタウン・インダストリアル・エstate テクノロジー・キャンパス

(72) 発明者 ソグリン、ミハイル

アイルランド共和国 15 モルハダート ダマスタウン・インダストリアル・エstate テクノロジー・キャンパス

審査官 成瀬 博之

(56) 参考文献 特開2008-310784 (JP, A)

特開2016-212838 (JP, A)

特開2015-060243 (JP, A)

(58) 調査した分野 (Int.Cl., DB名)

G 06 F 40 / 20 - 40 / 58

G 06 N 5 / 00 - 7 / 06

G 06 F 16 / 00 - 16 / 958

G 06 Q 10 / 00 - 99 / 00