

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2025年6月26日(26.06.2025)



(10) 国際公開番号

WO 2025/134331 A1

(51) 国際特許分類:
G06F 16/33 (2019.01)

(21) 国際出願番号: PCT/JP2023/046002

(22) 国際出願日: 2023年12月21日(21.12.2023)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(71) 出願人: 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 (JP).

(72) 発明者: 中西 弘毅 (NAKANISHI, Koki); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 (JP). 高木 郁子 (TAKAGI, Ikuko); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 (JP). 坂本 昌史 (SAKAMOTO, Masanobu); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ

内 (JP). 大石 晴夫 (OISHI, Haruo); 〒1808585 東京都武蔵野市緑町3丁目9-11 NTT 知的財産センタ内 (JP).

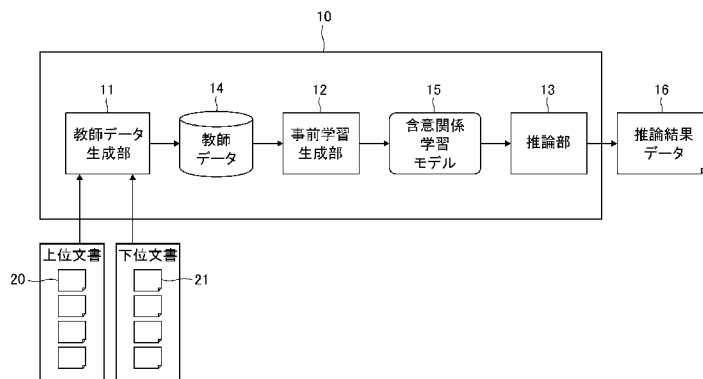
(74) 代理人: 弁理士法人鈴榮特許総合事務所 (SUZUYE & SUZUYE); 〒1050014 東京都港区芝三丁目23番1号 セレスティン芝三井ビルディング11階 (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(54) Title: NATURAL LANGUAGE PROCESSING DEVICE

(54) 発明の名称: 自然言語処理装置

[図1]



11... TEACHING DATA GENERATION UNIT
12... PRE-TRAINING GENERATION UNIT
13... INFERENCE UNIT
14... TEACHING DATA
15... IMPLICATIONAL RELATIONSHIP TRAINED MODEL
16... INFERENCE RESULTS DATA
20... UPPER-LEVEL DOCUMENTS
21... LOWER-LEVEL DOCUMENTS

(57) Abstract: This natural language processing device comprises a teaching data generation unit and an inference unit. The teaching data generation unit generates teaching data expressing correspondence relationships between lower-level documents and each of chapters, subchapters, and sections included in upper-level documents. The inference unit acquires a chapter and a subchapter of a first upper-level document, said chapter and subchapter including a section that serves as a search condition, and on the basis of a machine learning model that was trained using the teaching data, performs a search of the lower-level documents, in accordance with reference relationships between the lower-level documents and in the order of said chapter, said subchapter, and said section of the first upper-level document, and outputs a lower-level document corresponding to said section of the first upper-level document.

[続葉有]

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告 (条約第21条(3))

(57) 要約: 自然言語処理装置は、教師データ生成部と、推論部とを備える。教師データ生成部は、上位文書に含まれる章、節、及び項の各々と下位文書との対応関係を表す教師データを生成する。推論部は、検索条件としての項を含む第1上位文書の章及び節を取得し、前記教師データを用いて生成された機械学習モデルに基づいて、前記下位文書間の参照関係に従い前記第1上位文書の前記章、前記節、及び前記項の順で前記下位文書に対する検索を行い、前記第1上位文書の前記項に対応する下位文書を出力する。

明 細 書

発明の名称：自然言語処理装置

技術分野

[0001] この発明は、自然言語処理装置に関する。

背景技術

[0002] 複数の要件が記載された上位文書と、上位文書に記載された個々の要件を満たすための手段が記載された下位文書があるとき、上位文書に書かれた要件に対して、その要件を満たすための手段が書かれた下位文書の個所との対応関係を確認することとなる。

先行技術文献

非特許文献

[0003] 非特許文献1：Jacob Devlin Ming-Wei Chang Kenton Lee Kristina Toutanova: "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", Proceedings of NAACL HLT 2019, pages 4171-4186

非特許文献2：Pretrained Japanese BERT models released / 日本語 BERT モデル公開インターネット<URL:<https://www.nlp.ecei.tohoku.ac.jp/newsrelease/3284/>>

発明の概要

発明が解決しようとする課題

[0004] 上位文書に記載された要件に対応する箇所を下位文書から探す場合、文書中の文と文の対応づけを行う機械学習（例えば、非特許文献1、2参照）を用いる方法が考えられる。

[0005] しかし、機械学習を用いる場合、その教師データの作成において上位文書と下位文書との間に含意関係のラベル付けを行う際に、項目（あるいは、項目）ごとの対応関係をラベル付けしていたために、章、節の文脈を考慮することができないという課題がある。

[0006] この発明は、上記事情に着目してなされたもので、その目的は、要件が記

載された上位文書の章、節、及び項の文脈を考慮した要件を満たす下位文書の検索を行うことが可能な自然言語処理装置を提供することである。

課題を解決するための手段

[0007] この発明に係る一態様の自然言語処理装置は、教師データ生成部と、推論部とを備える。教師データ生成部は、上位文書に含まれる章、節、及び項の各々と下位文書との対応関係を表す教師データを生成する。推論部は、検索条件としての項を含む第1上位文書の章及び節を取得し、前記教師データを用いて生成された機械学習モデルに基づいて、前記下位文書間の参照関係に従い前記第1上位文書の前記章、前記節、及び前記項の順で前記下位文書に対する検索を行い、前記第1上位文書の前記項に対応する下位文書を出力する。

発明の効果

[0008] この発明によれば、要件が記載された上位文書の章、節、及び項の文脈を考慮した要件を満たす下位文書の検索を行うことが可能な自然言語処理装置が提供できる。

図面の簡単な説明

[0009] [図1]図1は、実施形態に係る自然言語処理装置の機能構成を示すブロック図である。

[図2]図2は、実施形態に係る上位文書あるいは下位文書における章、節、及び項の一例を示す図である。

[図3]図3は、実施形態に係る自然言語処理装置における自然言語処理を示すフローチャートである。

[図4]図4は、実施形態に係る教師データ生成部における教師データの生成処理を示すフローチャートである。

[図5]図5は、実施形態に係る教師データの生成処理の様子を示す図である。

[図6]図6は、実施形態に係る上位文書の章と下位文書との間の対応関係を表すマトリクスを示す図である。

[図7]図7は、実施形態に係る上位文書の節と下位文書との間の対応関係を表

すマトリクスを示す図である。

[図8]図8は、実施形態に係る上位文書の項と下位文書との間の対応関係を表すマトリクスを示す図である。

[図9]図9は、実施形態に係る推論部における推論結果データの生成処理を示すフローチャートである。

[図10]図10は、実施形態に係る推論結果データの生成処理の様子を示す図である。

[図11]図11は、実施形態に係る推論結果データの生成処理の様子を別の視点から示す図である。

[図12]図12は、実施形態に係る自然言語処理装置のハードウェア構成の一例を示す図である。

発明を実施するための形態

[0010] 以下、図面を参照してこの発明に係る実施形態について説明する。以下の説明において、同一の機能及び構成を有する構成要素については、共通する参照符号を付す。

[0011] 1. 実施形態の機能構成

図1を参照して、実施形態の自然言語処理装置10の機能構成について説明する。図1は、自然言語処理装置10の機能構成を示すブロック図である。

[0012] 実施形態の自然言語処理装置10は、上位文書の要件項目に対応（あるいは、適合）する下位文書を検索する際、要件項目が記載された上位文書の章、節、及び項の文脈を考慮しつつ下位文書を検索し抽出する装置である。

[0013] ここで、上位文書、下位文書、章、節、及び項について説明しておく。上位文書は、例えば、複数の要件が記載された文書である。下位文書は、例えば、上位文書に記載された個々の要件を満たすための手段が記載された文書である。

[0014] 上位文書あるいは下位文書は、章、節、及び項の少なくともいずれか1つを含む。章は1つ以上の節を含む。節は1つ以上の項を含む。例えば、章は

、上位文書あるいは下位文書に記載された文章全体をいくつかに分けた大きな区分を表す。節は、章をいくつかに分けた区分を表す。さらに、項は、節をいくつかに分けた小さな区分を表す。この明細書に記載される「章」は章が含む文あるいは文意を表し、「節」は節が含む文あるいは文意を表し、「項」は項が含む文あるいは文意を表すものとする。

[0015] 図2を参照して、上位文書あるいは下位文書における章、節、及び項の構成の一例について説明する。図2は、上位文書あるいは下位文書における章、節、及び項の一例を示す図である。上位文書20あるいは下位文書21は、章1及び章2を含む。章1は、節1、節2、及び節3を含む。節1は、項1、項2、及び項3を含む。節2は、項1及び項2を含む。節3は項1を含む。章2は、節1及び節2を含む。節1は、項1及び項2を含む。節2は項1を含む。なお、上位文書及び下位文書における章、節、及び項の構成及び数は上記に限るものではない。

[0016] 図1に示すように、自然言語処理装置10は、教師データ生成部11、事前学習生成部12、及び推論部13を備える。自然言語処理装置10は、さらに、データとして教師データ14、及び機械学習モデル（あるいは、深層学習モデル）として含意関係学習モデル15を備える。

[0017] 教師データ生成部11には、入力データとして、複数の上位文書20と複数の下位文書21が入力される。教師データ生成部11は、複数の上位文書20と複数の下位文書21とを用いて、教師データ14を生成する。教師データ生成部11は、上位文書20に含まれる章、節、及び項の各々と下位文書21との間の対応関係を判定し、対応関係の有無を示すラベルを生成する。すなわち、教師データ生成部11は、上位文書20の章、節、及び項の各々と下位文書21との間の含意関係を判定し、含意関係が成り立つか否かを示すラベルを生成する。具体的には、教師データ生成部11は、上位文書20に含まれる章、節、及び項の各々に記載された文（あるいは、文意）毎に、下位文書21に記載された文（あるいは、文意）との間にそれぞれ含意関係が成り立つか否かを示すラベルを生成する。含意関係が成り立つとは、例

えば、上位文書 20 の章が真であるなら下位文書 21 も真であるとき、上位文書 20 の章と下位文書 21 との間に含意関係が成り立つと判定される。

[0018] 事前学習生成部 12 は、教師データ生成部 11 により生成された教師データ 14 を入力とする機械学習アルゴリズムによって、上位文書 20 の章、節、及び項の各々と下位文書 21 との含意関係を学習し、含意関係を判定するための含意関係学習モデル 15 を生成する。含意関係学習モデル 15 は、上位文書の章、節、及び項の各々が下位文書を含意するか否かを推論すると共に、推論することを学習するモデルである。

[0019] 推論部 13 は、要件項目（あるいは、項）を含む上位文書の章及び節の情報を取得し、含意関係学習モデル 15 に基づいて、章、節、及び要件項目の順に、章、節、及び要件項目の各々が含意する下位文書 21 を抽出する。推論部 13 は、また含意関係学習モデル 15 あるいは他の自然言語処理モデルに従って、複数の下位文書 21 間の参照関係を生成する、言い換えると複数の下位文書 21 間にリンクを設定する。推論部 13 は、検索条件（あるいは、検索キー）としての要件項目を含む上位文書の章、節、及び要件項目の順に下位文書 21 を検索するとともに、下位文書 21 間の参照関係に従って下位文書 21 を絞り込むことにより、要件項目に対応（あるいは、適合）する下位文書を抽出する。そして、推論部 13 は、抽出された下位文書を含む推論結果データ 16 を出力する。

[0020] 事前学習生成部 12 が用いる機械学習アルゴリズム及び推論部 13 が用いる自然言語処理モデルには、例えば、非特許文献 1 や非特許文献 2 に記載された BERT (Bidirectional Encoder Representations from Transformers) が適用可能である。

[0021] 2. 実施形態の動作

図 3 を参照して、実施形態の自然言語処理装置 10 の動作について説明する。図 3 は、自然言語処理装置 10 が実行する自然言語処理を示すフローチャートである。

[0022] 先ず、教師データ生成部 11 は、複数の上位文書 20 及び複数の下位文書

21を受け取る（S1）。

[0023] 次に、教師データ生成部11は、受け取った上位文書20及び下位文書21を用いて、教師データ14を生成する（S2）。教師データ14の生成処理については後で詳述する。

[0024] 次に、事前学習生成部12は、教師データ生成部11により生成された教師データ14を用いて、含意関係学習モデル15を生成する（S3）。

[0025] 次に、推論部13は、検索条件としての要件項目を含む上位文書の章、節、及び要件項目、含意関係学習モデル15、並びに下位文書間の参照関係に基づいて、要件項目に対応する下位文書を検索し、検索により抽出された下位文書を含む推論結果データ16を出力する（S4）。推論結果データ16の生成処理については後で詳述する。

[0026] 2. 1 教師データの生成処理

次に、図4～図8を参照して、教師データ生成部11が実行する教師データ14の生成処理について説明する。図4は、教師データ生成部11における教師データ14の生成処理を示すフローチャートである。図5は、教師データ14の生成処理の様子を示す図である。図6、図7、及び図8は、教師データ生成部11によって生成される教師データを示す。詳述すると、図6は、上位文書の章と下位文書との間の対応関係を表すマトリクス（あるいは、テーブル）を示す。図7は、上位文書の節と下位文書との間の対応関係を表すマトリクス（あるいは、テーブル）を示す。さらに、図8は、上位文書の項と下位文書との間の対応関係を表すマトリクス（あるいは、テーブル）を示す。ここでは、複数の上位文書20_1、20_2、20_3、…、20_mと、複数の下位文書21_1、21_2、21_3、…、21_nが用意されているものとする。m、nは2以上の整数である。

[0027] 前述したように、教師データ生成部11は、上位文書20_1～20_mの章、節、及び項の各々に対して、下位文書21_1～21_nをそれぞれ対応付ける。すなわち、教師データ生成部11は、上位文書20_1～20_mの章、節、及び項の各々と下位文書21_1～21_nとの間のそれぞれの対応

関係を示すラベルを生成する。言い換えると、上位文書 20_1 ~ 20_m の章、節、及び項の各々と下位文書 21_1 ~ 21_n との間にそれぞれ含意関係が成り立つか否かを示すラベルを生成する。

[0028] 先ず、教師データ生成部 11 は、図 4 及び図 5 に示すように、上位文書 20_1 ~ 20_m の章と下位文書 21_1 ~ 21_n とを対応付ける (S 11) 。上位文書の章が下位文書を含意するか否かによって、上位文書の章と下位文書とが対応関係に有るか否かを示すラベルを生成する。すなわち、上位文書の章に記載された文（あるいは、文意）が、下位文書に記載された文（あるいは、文意）を含意する場合、上位文書の章と下位文書との間に含意関係が成り立つことを示すラベルを生成する。逆に、上位文書の章に記載された文が、下位文書に記載された文を含意しない場合、上位文書の章と下位文書との間に含意関係が成り立たないことを示すラベルを生成する。

[0029] 具体的には、教師データ生成部 11 は、上位文書 20_1 の章 1 と下位文書 21_1 とが対応関係に有るか否かを判定する。すなわち、上位文書 20_1 の章 1 が下位文書 21_1 を含意するか否かを判定する。図 6 に示すように、教師データ生成部 11 は、上位文書 20_1 の章 1 が下位文書 21_1 を含意する場合、上位文書 20_1 の章 1 と下位文書 21_1 とが対応関係に有ることを示すラベル 1 を生成する。上位文書 20_1 の章 1 が下位文書 21_1 を含意しない場合、上位文書 20_1 の章 1 と下位文書 21_1 とが対応関係に無いことを示すラベル 0 を生成する。

[0030] 次に、教師データ生成部 11 は、上位文書 20_1 の章 1 と下位文書 21_2 とが対応関係に有るか否かを判定する。すなわち、上位文書 20_1 の章 1 が下位文書 21_2 を含意するか否かを判定する。図 6 に示すように、教師データ生成部 11 は、上位文書 20_1 の章 1 が下位文書 21_2 を含意する場合、上位文書 20_1 の章 1 と下位文書 21_2 とが対応関係に有ることを示すラベル 1 を生成する。上位文書 20_1 の章 1 が下位文書 21_2 を含意しない場合、上位文書 20_1 の章 1 と下位文書 21_2 とが対応関係に無いことを示すラベル 0 を生成する。

- [0031] 同様に、教師データ生成部11は、上位文書20_1の章1と下位文書21_3～21_nの各々とが対応関係に有るか否かを判定する。すなわち、上位文書20_1の章1が下位文書21_3～21_nの各々を含意するか否かを判定する。図6に示すように、教師データ生成部11は、上位文書20_1の章1が下位文書21_3～21_nの各々を含意する場合、上位文書20_1の章1と下位文書21_3～21_nの各々とが対応関係に有ることを示すラベル1を生成する。上位文書20_1の章1が下位文書21_3～21_nの各々を含意しない場合、上位文書20_1の章1と下位文書21_3～21_nの各々とが対応関係に無いことを示すラベル0を生成する。
- [0032] 以下同様に、教師データ生成部11は、上位文書20_1の章2と下位文書21_1～21_nの各々とが対応関係に有るか否かを判定する。すなわち、上位文書20_1の章2が下位文書21_1～21_nの各々を含意するか否かを判定する。図6に示すように、教師データ生成部11は、上位文書20_1の章2が下位文書21_1～21_nの各々を含意する場合、上位文書20_1の章2と下位文書21_1～21_nの各々とが対応関係に有ることを示すラベル1を生成する。上位文書20_1の章2が下位文書21_1～21_nの各々を含意しない場合、上位文書20_1の章2と下位文書21_1～21_nの各々とが対応関係に無いことを示すラベル0を生成する。
- [0033] 上位文書20_1の章3以降の各々と下位文書21_1～21_nの各々についても上述と同様である。教師データ生成部11は、上位文書20_1の章3以降の各々と下位文書21_1～21_nの各々とが対応関係に有るか否かを判定する。教師データ生成部11は、上位文書20_1の章3以降の各々が下位文書21_1～21_nの各々を含意する場合、ラベル1をそれぞれ生成する。上位文書20_1の章3以降の各々が下位文書21_1～21_nの各々をそれぞれ含意しない場合、ラベル0をそれぞれ生成する。
- [0034] 上位文書20_2の章の各々と下位文書21_1～21_nの各々についても上述と同様である。教師データ生成部11は、上位文書20_2の章の各々と下位文書21_1～21_nの各々とが対応関係に有るか否かを判定する。

教師データ生成部 11 は、上位文書 20_2 の章の各々が下位文書 21_1 ~ 21_n の各々を含意する場合、ラベル 1 をそれぞれ生成する。上位文書 20_2 の章の各々が下位文書 21_1 ~ 21_n の各々を含意しない場合、ラベル 0 をそれぞれ生成する。

[0035] 上位文書 20_3 ~ 20_m の章の各々と下位文書 21_1 ~ 21_n の各々についても上述と同様である。教師データ生成部 11 は、上位文書 20_3 ~ 20_m の章の各々と下位文書 21_1 ~ 21_n の各々とが対応関係に有るか否かを判定し、ラベル 1 あるいはラベル 0 をそれぞれ生成する。

[0036] 次に、教師データ生成部 11 は、図 4 及び図 5 に示すように、上位文書 20_1 ~ 20_m の節と下位文書 21_1 ~ 21_n とを対応付ける (S12)。上位文書の節が下位文書を含意するか否かによって、上位文書の節と下位文書とが対応関係に有るか否かを示すラベルを生成する。すなわち、上位文書の節に記載された文（あるいは、文意）が、下位文書に記載された文（あるいは、文意）を含意する場合、上位文書の節と下位文書との間に含意関係が成り立つことを示すラベルを生成する。逆に、上位文書の節に記載された文が、下位文書に記載された文を含意しない場合、上位文書の節と下位文書との間に含意関係が成り立たないことを示すラベルを生成する。

[0037] 具体的には、教師データ生成部 11 は、上位文書 20_1 章 1 の節 1 と下位文書 21_1 とが対応関係に有るか否かを判定する。すなわち、上位文書 20_1 章 1 の節 1 が下位文書 21_1 を含意するか否かを判定する。図 7 に示すように、教師データ生成部 11 は、上位文書 20_1 章 1 の節 1 が下位文書 21_1 を含意する場合、上位文書 20_1 章 1 の節 1 と下位文書 21_1 とが対応関係に有ることを示すラベル 1 を生成する。上位文書 20_1 章 1 の節 1 が下位文書 21_1 を含意しない場合、上位文書 20_1 章 1 の節 1 と下位文書 21_1 とが対応関係に無いことを示すラベル 0 を生成する。

[0038] 次に、教師データ生成部 11 は、上位文書 20_1 章 1 の節 1 と下位文書 21_2 とが対応関係に有るか否かを判定する。すなわち、上位文書 20_1 章 1 の節 1 が下位文書 21_2 を含意するか否かを判定する。図 7 に示すように

、教師データ生成部 11 は、上位文書 20_1 章 1 の節 1 が下位文書 21_2 を含意する場合、上位文書 20_1 章 1 の節 1 と下位文書 21_2 とが対応関係に有ることを示すラベル 1 を生成する。上位文書 20_1 章 1 の節 1 が下位文書 21_2 を含意しない場合、上位文書 20_1 章 1 の節 1 と下位文書 21_2 とが対応関係に無いことを示すラベル 0 を生成する。

[0039] 同様に、教師データ生成部 11 は、上位文書 20_1 章 1 の節 1 と下位文書 21_3～21_n の各々とが対応関係に有るか否かを判定する。すなわち、上位文書 20_1 章 1 の節 1 が下位文書 21_3～21_n の各々を含意するか否かを判定する。図 7 に示すように、教師データ生成部 11 は、上位文書 20_1 章 1 の節 1 が下位文書 21_3～21_n の各々を含意する場合、上位文書 20_1 章 1 の節 1 と下位文書 21_3～21_n の各々とが対応関係に有ることを示すラベル 1 を生成する。上位文書 20_1 章 1 の節 1 が下位文書 21_3～21_n の各々を含意しない場合、上位文書 20_1 章 1 の節 1 と下位文書 21_3～21_n の各々とが対応関係に無いことを示すラベル 0 を生成する。

[0040] 以下同様に、教師データ生成部 11 は、上位文書 20_1 章 1 の節 2 と下位文書 21_1～21_n の各々とが対応関係に有るか否かを判定する。すなわち、上位文書 20_1 章 1 の節 2 が下位文書 21_1～21_n の各々を含意するか否かを判定する。図 7 に示すように、教師データ生成部 11 は、上位文書 20_1 章 1 の節 2 が下位文書 21_1～21_n の各々を含意する場合、上位文書 20_1 章 1 の節 2 と下位文書 21_1～21_n の各々とが対応関係に有ることを示すラベル 1 を生成する。上位文書 20_1 章 1 の節 2 が下位文書 21_1～21_n の各々を含意しない場合、上位文書 20_1 章 1 の節 2 と下位文書 21_1～21_n の各々とが対応関係に無いことを示すラベル 0 を生成する。

[0041] 上位文書 20_1 章 1 の節 3 以降の各々と下位文書 21_1～21_n の各々についても上述と同様である。教師データ生成部 11 は、上位文書 20_1 章 1 の節 3 以降の各々と下位文書 21_1～21_n の各々とが対応関係に有

るか否かを判定する。教師データ生成部 11 は、上位文書 20_1 章 1 の節 3 以降の各々が下位文書 21_1 ~ 21_n の各々を含意する場合、ラベル 1 をそれぞれ生成する。上位文書 20_1 章 1 の節 3 以降の各々が下位文書 21_1 ~ 21_n の各々をそれぞれ含意しない場合、ラベル 0 をそれぞれ生成する。

[0042] 上位文書 20_1 章 2 の節の各々と下位文書 21_1 ~ 21_n の各々についても上述と同様である。教師データ生成部 11 は、上位文書 20_1 章 2 の節の各々と下位文書 21_1 ~ 21_n の各々が対応関係に有るか否かを判定する。教師データ生成部 11 は、上位文書 20_1 章 2 の節の各々が下位文書 21_1 ~ 21_n の各々を含意する場合、ラベル 1 をそれぞれ生成する。上位文書 20_1 章 2 の節の各々が下位文書 21_1 ~ 21_n の各々を含意しない場合、ラベル 0 をそれぞれ生成する。

[0043] 上位文書 20_1 ~ 20_m のその他の節の各々と下位文書 21_1 ~ 21_n の各々についても上述と同様であり、記載を省略する。

[0044] 次に、教師データ生成部 11 は、図 4 及び図 5 に示すように、上位文書 20_1 ~ 20_m の項と下位文書 21_1 ~ 21_n とを対応付ける (S13)。上位文書の項が下位文書を含意するか否かによって、上位文書の項と下位文書とが対応関係に有るか否かを示すラベルを生成する。すなわち、上位文書の項に記載された文（あるいは、文意）が、下位文書に記載された文（あるいは、文意）を含意する場合、上位文書の項と下位文書との間に含意関係が成り立つことを示すラベルを生成する。逆に、上位文書の項に記載された文が、下位文書に記載された文を含意しない場合、上位文書の項と下位文書との間に含意関係が成り立たないことを示すラベルを生成する。

[0045] 具体的には、教師データ生成部 11 は、上位文書 20_1 章 1 節 1 の項 1 と下位文書 21_1 とが対応関係に有るか否かを判定する。すなわち、上位文書 20_1 章 1 節 1 の項 1 が下位文書 21_1 を含意するか否かを判定する。図 8 に示すように、教師データ生成部 11 は、上位文書 20_1 章 1 節 1 の項 1 が下位文書 21_1 を含意する場合、上位文書 20_1 章 1 節 1 の項 1 と下位

文書 2 1 _ 1 とが対応関係に有ることを示すラベル 1 を生成する。上位文書 2 0 _ 1 章 1 節 1 の項 1 が下位文書 2 1 _ 1 を含意しない場合、上位文書 2 0 _ 1 章 1 節 1 の項 1 と下位文書 2 1 _ 1 とが対応関係に無いことを示すラベル 0 を生成する。

[0046] 次に、教師データ生成部 1 1 は、上位文書 2 0 _ 1 章 1 節 1 の項 1 と下位文書 2 1 _ 2 とが対応関係に有るか否かを判定する。すなわち、上位文書 2 0 _ 1 章 1 節 1 の項 1 が下位文書 2 1 _ 2 を含意するか否かを判定する。図 8 に示すように、教師データ生成部 1 1 は、上位文書 2 0 _ 1 章 1 節 1 の項 1 が下位文書 2 1 _ 2 を含意する場合、上位文書 2 0 _ 1 章 1 節 1 の項 1 と下位文書 2 1 _ 2 とが対応関係に有ることを示すラベル 1 を生成する。上位文書 2 0 _ 1 章 1 節 1 の項 1 が下位文書 2 1 _ 2 を含意しない場合、上位文書 2 0 _ 1 章 1 節 1 の項 1 と下位文書 2 1 _ 2 とが対応関係に無いことを示すラベル 0 を生成する。

[0047] 同様に、教師データ生成部 1 1 は、上位文書 2 0 _ 1 章 1 節 1 の項 1 と下位文書 2 1 _ 3 ~ 2 1 _ n の各々とが対応関係に有るか否かを判定する。すなわち、上位文書 2 0 _ 1 章 1 節 1 の項 1 が下位文書 2 1 _ 3 ~ 2 1 _ n の各々を含意するか否かを判定する。図 8 に示すように、教師データ生成部 1 1 は、上位文書 2 0 _ 1 章 1 節 1 の項 1 が下位文書 2 1 _ 3 ~ 2 1 _ n の各々を含意する場合、上位文書 2 0 _ 1 章 1 節 1 の項 1 と下位文書 2 1 _ 3 ~ 2 1 _ n の各々とが対応関係に有ることを示すラベル 1 を生成する。上位文書 2 0 _ 1 章 1 節 1 の項 1 が下位文書 2 1 _ 3 ~ 2 1 _ n の各々を含意しない場合、上位文書 2 0 _ 1 章 1 節 1 の項 1 と下位文書 2 1 _ 3 ~ 2 1 _ n の各々とが対応関係に無いことを示すラベル 0 を生成する。

[0048] 以下同様に、教師データ生成部 1 1 は、上位文書 2 0 _ 1 章 1 節 1 の項 2 と下位文書 2 1 _ 1 ~ 2 1 _ n の各々とが対応関係に有るか否かを判定する。すなわち、上位文書 2 0 _ 1 章 1 節 1 の項 2 が下位文書 2 1 _ 1 ~ 2 1 _ n の各々を含意するか否かを判定する。図 8 に示すように、教師データ生成部 1 1 は、上位文書 2 0 _ 1 章 1 節 1 の項 2 が下位文書 2 1 _ 1 ~ 2 1 _ n の各々を含意

する場合、上位文書 20_1 章 1 節 1 の項 2 と下位文書 21_1 ~ 21_n の各々とが対応関係に有ることを示すラベル 1 を生成する。上位文書 20_1 章 1 節 1 の項 2 が下位文書 21_1 ~ 21_n の各々を含意しない場合、上位文書 20_1 章 1 節 1 の項 2 と下位文書 21_1 ~ 21_n の各々とが対応関係に無いことを示すラベル 0 を生成する。

[0049] 上位文書 20_1 章 1 節 1 の項 3 以降の各々と下位文書 21_1 ~ 21_n の各々についても上述と同様である。教師データ生成部 11 は、上位文書 20_1 章 1 節 1 の項 3 以降の各々と下位文書 21_1 ~ 21_n の各々とが対応関係に有るか否かを判定する。教師データ生成部 11 は、上位文書 20_1 章 1 節 1 の項 3 以降の各々が下位文書 21_1 ~ 21_n の各々を含意する場合、ラベル 1 をそれぞれ生成する。上位文書 20_1 章 1 節 1 の項 3 以降の各々が下位文書 21_1 ~ 21_n の各々をそれぞれ含意しない場合、ラベル 0 をそれぞれ生成する。

[0050] 上位文書 20_1 ~ 20_m のその他の項の各々と下位文書 21_1 ~ 21_n の各々についても上述と同様であり、記載を省略する。

[0051] 2. 2 推論結果データの生成処理

次に、図 9 ~ 図 10 を参照して、推論部 13 が実行する推論結果データ 16 の生成処理（推論処理）について説明する。図 9 は、推論部 13 における推論結果データ 16 の生成処理を示すフローチャートである。図 10 は、推論結果データ 16 の生成処理の様子を示す図である。なお、上位文書 20a は章 A を含み、章 A は節 B を含み、節 B は要件項目（あるいは、項）C を含むものとする。以下に、検索条件（あるいは、検索キー）としての要件項目 C を取得し、要件項目 C に対応する下位文書を抽出する場合の処理を説明する。

[0052] 先ず、推論部 13 は、図 9 及び図 10 に示すように、要件項目 C を含む上位文書 20a の章 A を取得し、取得した章 A が含意する下位文書 21a を抽出する（S21）。すなわち、推論部 13 は、要件項目 C を含む上位文書 20a の章 A に対応付けられた下位文書 21a を抽出する。

[0053] 次に、推論部 13 は、要件項目 C を含む上位文書 20 a の節 B を取得する。さらに、推論部 13 は、ステップ S 21 で抽出された下位文書 21 a から参照される下位文書の中から、言い換えると、下位文書 21 a にリンクが設定された下位文書の中から、取得した節 B が含意する下位文書 21 b を抽出する (S 22)。すなわち、推論部 13 は、要件項目 C を含む上位文書 20 a の節 B に対応付けられた下位文書 21 b を抽出する。

[0054] 次に、推論部 13 は、ステップ S 22 で抽出された下位文書 21 b から参照される下位文書の中から、言い換えると、下位文書 21 b にリンクが設定された下位文書の中から、要件項目 C が含意する下位文書 21 c を抽出する (S 23)。すなわち、推論部 13 は、要件項目 C に対応付けられた下位文書 21 c を抽出する。そして、推論部 13 は、抽出された下位文書 21 c を含む推論結果データ 16 を出力する。以上により、推論部 13 による推論結果データ 16 の生成処理 (推論処理) が終了する。

[0055] なお、上述した推論結果データ 16 の生成処理では、要件項目 C を含む上位文書 20 a の章 A を取得し、取得した章 A が含意する下位文書 21 a を抽出し、続いて、要件項目 C を含む上位文書 20 a の節 B を取得し、取得した節 B が含意する下位文書 21 a を抽出したが、これに限るわけではない。先に、要件項目 C を含む上位文書 20 a の章 A 及び節 B を取得し、取得した章 A が含意する下位文書 21 a を抽出し、続いて、取得した節 B が含意する下位文書 21 b を抽出するようにしてもよい。

[0056] 次に、図 11 を参照して、推論部 13 が実行する推論結果データ 16 の生成処理 (推論処理) について説明する。図 11 は、推論結果データ 16 の生成処理の様子を別の視点から示す図である。なお、上位文書 20 a は条件 1 を含み、条件 1 は条件 2 及び条件 3 を含む。条件 2 は要件 1 を含み、条件 3 は要件 2 を含む。また、条件 1 は上位文書 20 a の章 A に相当し、条件 2 は節 B に相当し、要件 1 は要件項目 C に相当する。条件 3 は他の節に相当し、要件 2 は他の要件項目に相当する。以下に、検索条件としての要件 1 を含む上位文書 20 a の条件 1 から、要件 1 に対応する下位文書を抽出する場合の

処理を説明する。

[0057] 先ず、推論部 13 は、図 11 に示すように、要件 1 を含む上位文書 20 a の条件 1 を取得し、取得した条件 1 が含意する下位文書 21 a を抽出する。すなわち、推論部 13 は、要件 1 を含む上位文書 20 a の条件 1 に対応付けられた下位文書 21 a を抽出する。

[0058] 次に、推論部 13 は、要件 1 を含む上位文書 20 a の条件 2 を取得する。さらに、推論部 13 は、取得した下位文書 21 a から参照される下位文書の中から、言い換えると、下位文書 21 a にリンクが設定された下位文書の中から、取得した条件 2 が含意する下位文書 21 b を抽出する。すなわち、推論部 13 は、要件 1 を含む上位文書 20 a の条件 2 に対応付けられた下位文書 21 b を抽出する。

[0059] 次に、推論部 13 は、抽出された下位文書 21 b から参照される下位文書の中から、言い換えると、下位文書 21 b にリンクが設定された下位文書の中から、要件 1 が含意する下位文書 21 c を抽出する。すなわち、推論部 13 は、要件 1 に対応付けられた下位文書 21 c を抽出する。そして、推論部 13 は、抽出された下位文書 21 c を含む推論結果データ 16 を出力する。以上により、推論部 13 による推論結果データ 16 の生成処理（推論処理）が終了する。

[0060] 以上述べたように実施形態では、要件項目 C（あるいは、要件 1）が記載された上位文書の章、節、及び項の順に下位文書が検索され、要件項目 C に適合する下位文書が抽出され出力される。

[0061] 3. 実施形態のハードウェア構成

次に、図 12 を参照して、実施形態の自然言語処理装置 10 のハードウェア構成について説明する。ここでは、自然言語処理装置 10 が、コンピュータ 30 で構成される例について説明する。

[0062] 図 12 は、自然言語処理装置 10 のハードウェア構成の一例を示す図である。自然言語処理装置 10（即ち、コンピュータ 30）は、プロセッサ 31、ROM（Read Only Memory）32、RAM（Random Access Memory）33

、補助記憶装置 34、及び入出力インタフェース 35 を有する。

プロセッサ 31 と ROM 32 と RAM 33 と補助記憶装置 34 と入出力インタフェース 35 は、バス 36 を介して互いに電氣的に接続されており、バス 36 を介して互いにデータ及び信号のやり取りが可能である。

[0063] プロセッサ 31 は、例えば、CPU (Central Processing Unit) や GPU (Graphical Processing Unit) 等を含む汎用ハードウェアプロセッサで構成される。プロセッサ 31 は、ROM 32、RAM 33、補助記憶装置 34、及び入出力インタフェース 35 の全体を制御する。

[0064] ROM 32 は、主記憶装置の一部を構成する不揮発性メモリである。ROM 32 は、自然言語処理装置 10 の起動時に必要な起動プログラムを非一時的に記憶している。自然言語処理装置 10 は、プロセッサ 31 が ROM 32 内のプログラムを実行することにより起動する。ROM 32 は、例えば、EPROM (Erasable Programmable Read Only Memory) で構成され、起動プログラムに加えて、起動時の諸設定を記憶している。

[0065] RAM 33 は、主記憶装置の一部を構成する揮発性メモリである。RAM 33 は、プロセッサ 31 の処理に必要なプログラムとプログラムの実行に必要なデータを一時的に記憶する。プロセッサ 31 は、RAM 33 内のプログラムを実行することにより、RAM 33 内のデータを演算し、演算結果を RAM 33 に記憶させる。

[0066] 補助記憶装置 34 は、HDD (Hard Disk Drive) や SSD (Solid State Drive) 等の不揮発性メモリで構成される。補助記憶装置 34 は、プロセッサ 31 が実行するプログラムとプログラムの実行に必要なデータを非一時的に記憶している。プロセッサ 31 は、補助記憶装置 34 内のプログラムとデータを RAM 33 内に読み込み、プログラムを実行することにより各種機能を実行する。

[0067] 入出力インタフェース 35 は、外部の入力装置 41 や出力装置 42 等と接続され、入力装置 41 からの情報の入力、出力装置 42 への情報の出力を可能にする。例えば、入出力インタフェース 35 は、有線のインタフェースで

あっても、無線のインタフェースであってもよい。有線のインタフェースは、装置が接続されるポート等を含む。無線のインタフェースは、Bluetooth（登録商標）、WiFi（登録商標）等を含む。

入力装置４１は、キーボード、マウス、タッチパネル、受信装置、ディスクドライブ等を含み得る。入力装置４１は、これに限らず、他の任意の入力機器を含み得る。出力装置４２は、ディスプレイ、送信装置、ディスクドライブ等を含み得る。出力装置４２は、これに限らず、他の任意の出力機器を含み得る。入力装置４１と出力装置４２は、両者の機能を有する入出力装置４３で構成されてもよい。

[0068] 入力データ、例えば、上位文書及び下位文書に関するデータ、及び検索対象としての上位文書の要件項目等のデータは、入力装置４１を介して、教師データ生成部１１、あるいは推論部１３に入力される。

[0069] 補助記憶装置３４に非一時的に記憶されるプログラムは、例えば、プログラムを非一時的に記録したコンピュータ３０で読み取り可能な記録媒体４４を介して、コンピュータ３０に提供される。そのような記録媒体４４は、非一時的コンピュータ読取可能記録媒体と呼ばれる。非一時的コンピュータ読取可能記録媒体は、フレキシブルディスク、光ディスク（ＣＤ－ＲＯＭ、ＣＤ－Ｒ、ＤＶＤ－ＲＯＭ、ＤＶＤ－Ｒ等）、光磁気ディスク（ＭＯ等）等のディスクや、半導体メモリ等を含む。

[0070] 補助記憶装置３４に非一時的に記憶されるプログラムは、記録媒体４４がディスクである場合には、入力装置４１であるディスクドライブと入出力インタフェース３５を介して、記録媒体４４が半導体メモリある場合には、入出力インタフェース３５であるポートを介して、補助記憶装置３４に読み込まれ非一時的に記憶される。また、プログラムは、ネットワーク上のサーバに格納されており、サーバからダウンロードされ、補助記憶装置３４に非一時的に記憶されてもよい。

[0071] プロセッサ３１は、コンピュータ３０の起動時に、ＲＯＭ３２内のプログラムを実行し、ＲＡＭ３３内にＯＳを読み込み起動する。プロセッサ３１は

、OSによる制御の下、指示入力や外部機器の接続等を監視する。また、プロセッサ31は、OSによる制御の下、RAM33内にプログラムエリアとデータエリアを設定する。プロセッサ31は、自然言語処理装置10の起動の指示入力に対して、補助記憶装置34から自然言語処理プログラムをRAM33のプログラムエリアに読み込むとともに、自然言語処理プログラムの実行に必要なデータを補助記憶装置34からRAM33のデータエリアに読み込む。プロセッサ31は、自然言語処理プログラムに従ってデータエリアのデータを演算し、演算結果をデータエリアに書き込む。このような動作によって、プロセッサ31とRAM33と補助記憶装置34と入出力インタフェース35とバス36は共働して、自然言語処理装置10の構成要素、すなわち教師データ生成部11、事前学習生成部12、及び推論部13の少なくとも一部の機能を実行する。

[0072] 4. 実施形態の効果等

この発明に係る実施形態によれば、要件が記載された上位文書の章、節、及び項の文脈を考慮した要件を満たす下位文書の検索を行うことが可能な自然言語処理装置が提供できる。

[0073] 実施形態の構成では、教師データ生成部11により、上位文書の章と下位文書との対応関係を示すラベル生成、上位文書の節と下位文書との対応関係を示すラベル生成、及び上位文書の項と下位文書との対応関係を示すラベル生成が実行されて、教師データ14が生成される。次に、事前学習生成部12により、教師データ14を入力データとして、機械学習モデル（例えば、BERT）を用いて、上位文書の章、節、及び項の各々と下位文書との間の含意関係を判定するための含意関係学習モデル15が生成される。そして、推論部13により、検索条件としての要件項目を含む上位文書の章及び節の情報を取得し、教師データ14を用いて生成された含意関係学習モデル15に基づいて、下位文書間の参照関係に従い上位文書の章、節、及び要件項目の順で前記下位文書を検索し絞り込む。これにより、要件項目を含む上位文書の章、節、及び要件項目の文脈を考慮した要件項目に対応する下位文書が

出力できる。

[0074] 言い換えると、実施形態の構成では、上位文書の章、節、及び項の各々と下位文書との対応関係を示すラベルを付与した教師データ 14 を作成し、上位文書の章、節、及び項の各々と下位文書との対応づけを推論する含意関係学習モデル 15 を生成する。推論時には、探索したい上位文書の要件項目を含む章と節の情報を取得し、下位文書間の参照関係を考慮しつつ章、節、及び要件項目の順に下位文書を検索し絞り込む。これにより、章、節、及び要件項目の文脈を考慮した要件項目を満たす下位文書の検索が可能である。

[0075] 上記実施形態で説明した機能ブロックは、ハードウェア、コンピュータソフトウェアのいずれかまたは両者を組み合わせたものとして実現することができる。機能ブロックが上述の例のように区別されていることは必須ではない。例えば、一部の機能が例示の機能ブロックとは別の機能ブロックによって実行されてもよい。さらに、例示の機能ブロックがさらに細かい機能サブブロックに分割されていてもよい。また、上記実施形態で説明したフローチャートは、その処理の順番を可能な限り入れ替えることができる。

[0076] なお、この発明は上記実施形態に限定されるものではない。例えば、この発明の要旨を逸脱しない範囲で種々変形して実施可能である。

[0077] 要するにこの発明は、上記実施形態そのままに限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で構成要素を変形して具体化できる。また、上記実施形態に開示されている複数の構成要素の適宜な組み合わせにより種々の発明を形成できる。例えば、実施形態に示される全構成要素から幾つかの構成要素を削除してもよい。さらに、異なる実施形態に亘る構成要素を適宜組み合わせてもよい。

符号の説明

- [0078] 10…自然言語処理装置
11…教師データ生成部
12…事前学習生成部
13…推論部

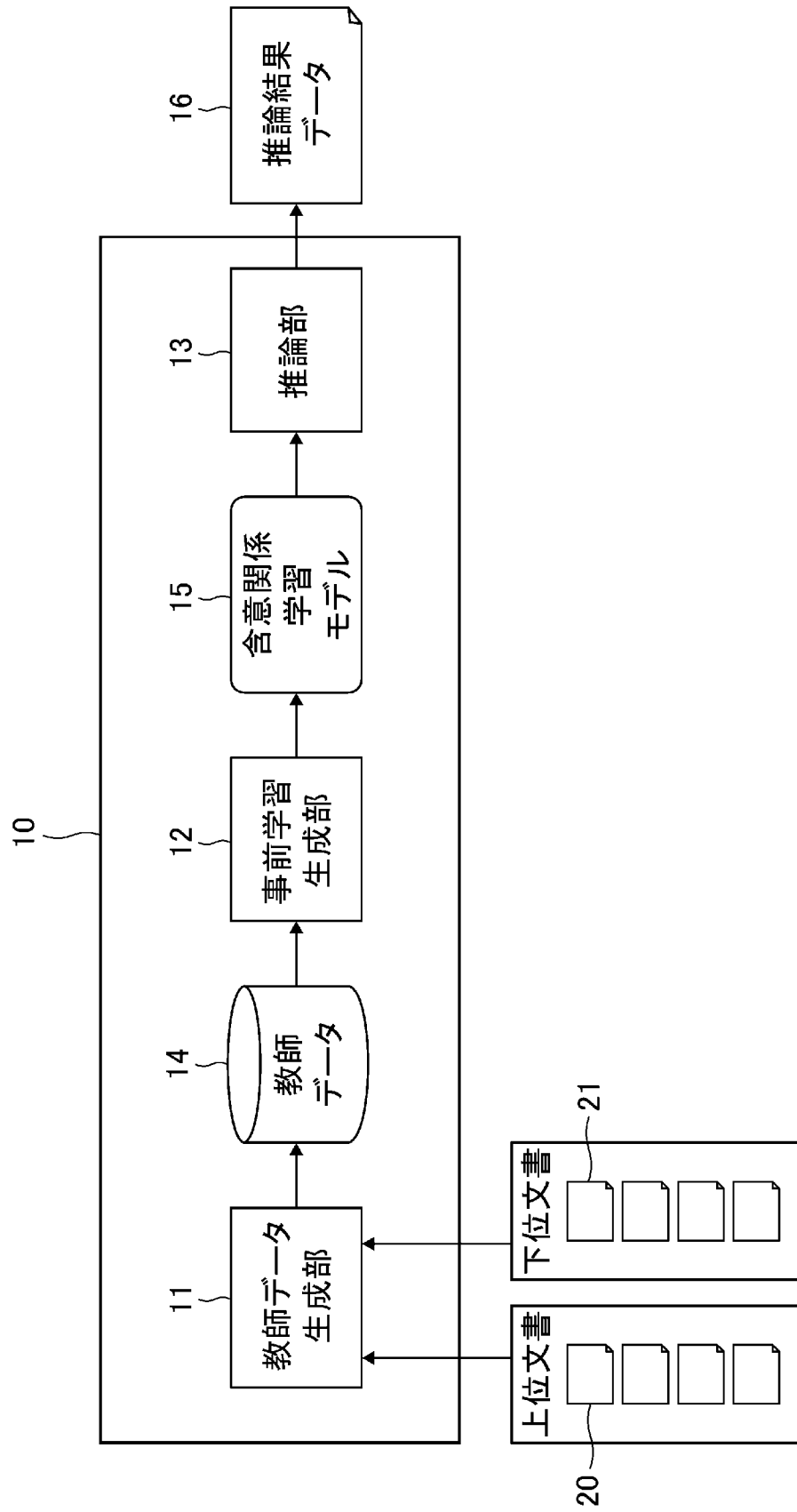
- 1 4 …教師データ
- 1 5 …含意関係学習モデル
- 1 6 …推論結果データ
- 2 0 …上位文書
- 2 0 a …上位文書
- 2 1 …下位文書
- 2 1 a …下位文書
- 2 1 b …下位文書
- 2 1 c …下位文書
- 3 0 …コンピュータ
- 3 1 …プロセッサ
- 3 2 …ROM
- 3 3 …RAM
- 3 4 …補助記憶装置
- 3 5 …入出力インタフェース
- 3 6 …バス
- 4 1 …入力装置
- 4 2 …出力装置
- 4 3 …入出力装置
- 4 4 …記録媒体。

請求の範囲

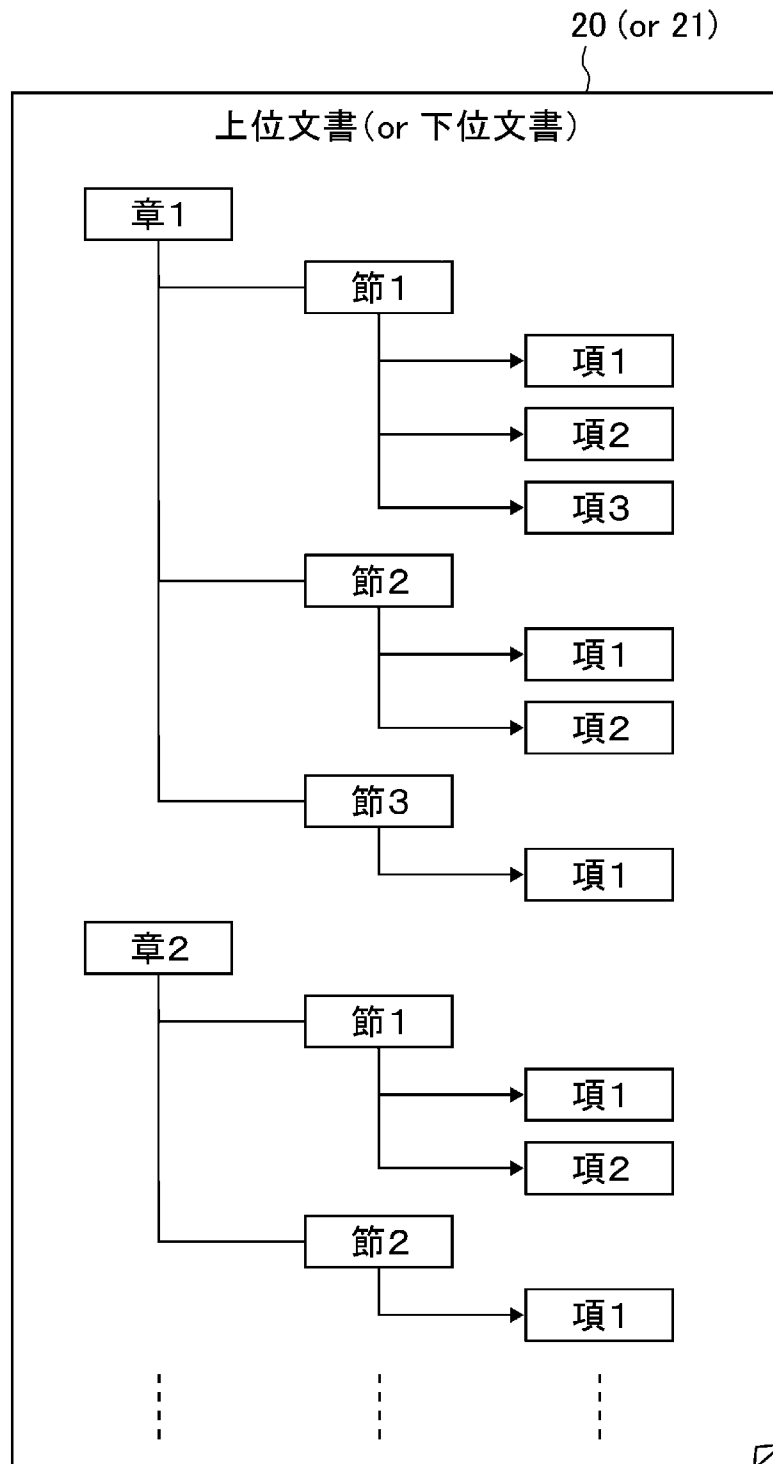
- [請求項1] 上位文書に含まれる章、節、及び項の各々と下位文書との対応関係を表す教師データを生成する教師データ生成部と、
- 検索条件としての項を含む第1上位文書の章及び節を取得し、前記教師データを用いて生成された機械学習モデルに基づいて、前記下位文書間の参照関係に従い前記第1上位文書の前記章、前記節、及び前記項の順で前記下位文書に対する検索を行い、前記第1上位文書の前記項に対応する下位文書を出力する推論部と、
- を備える自然言語処理装置。
- [請求項2] 前記教師データ生成部は、前記教師データとして、
- 前記上位文書の章と前記下位文書との間の含意関係を表すマトリクスと、前記上位文書の節と前記下位文書との間の含意関係を表すマトリクスと、前記上位文書の項と前記下位文書との間の含意関係を表すマトリクスと、を生成する、
- 請求項1に記載の自然言語処理装置。
- [請求項3] 前記推論部は、
- 前記下位文書の中から、前記第1上位文書の前記章が含意する第1下位文書を抽出し、
- 前記第1下位文書から参照される下位文書の中から、前記第1上位文書の前記節が含意する第2下位文書を抽出し、
- 前記第2下位文書から参照される下位文書の中から、前記第1上位文書の前記項が含意する第3下位文書を抽出し、抽出された前記第3下位文書を出力する、
- 請求項1に記載の自然言語処理装置。
- [請求項4] 前記機械学習モデルとして、前記教師データを用いて前記上位文書の章、節、及び項の各々と前記下位文書との間の含意関係を判定するための含意関係学習モデルを生成する事前学習生成部をさらに備える、

請求項 1 に記載の自然言語処理装置。

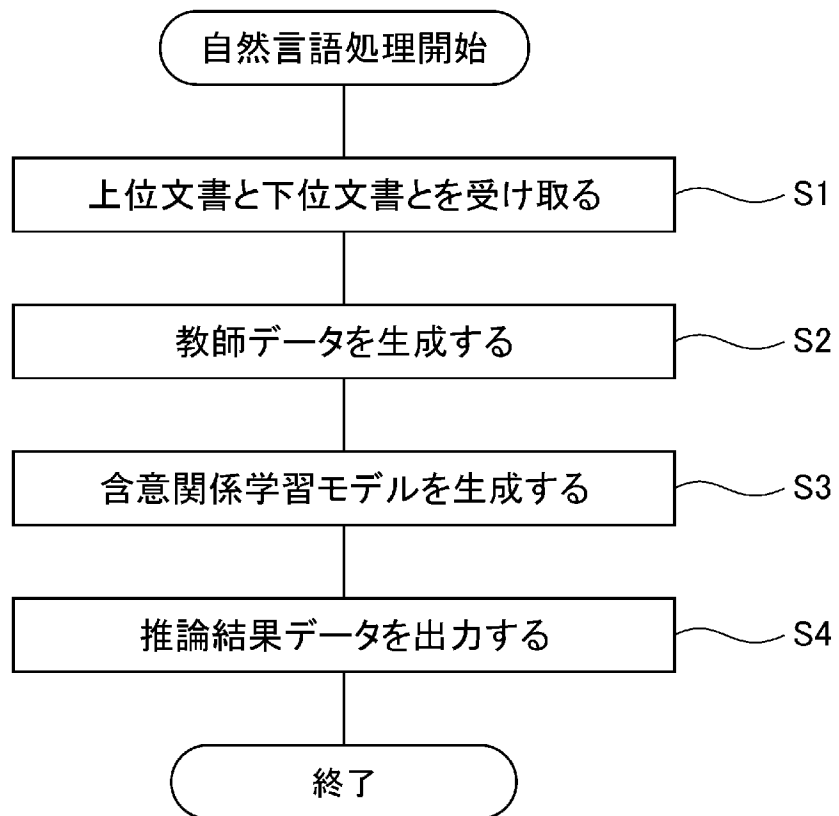
[図1]



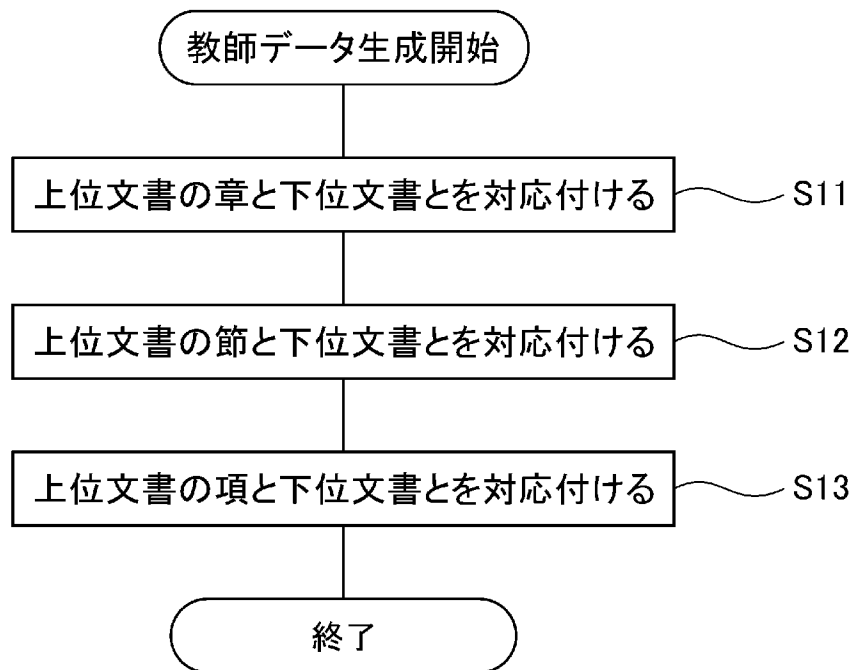
[図2]



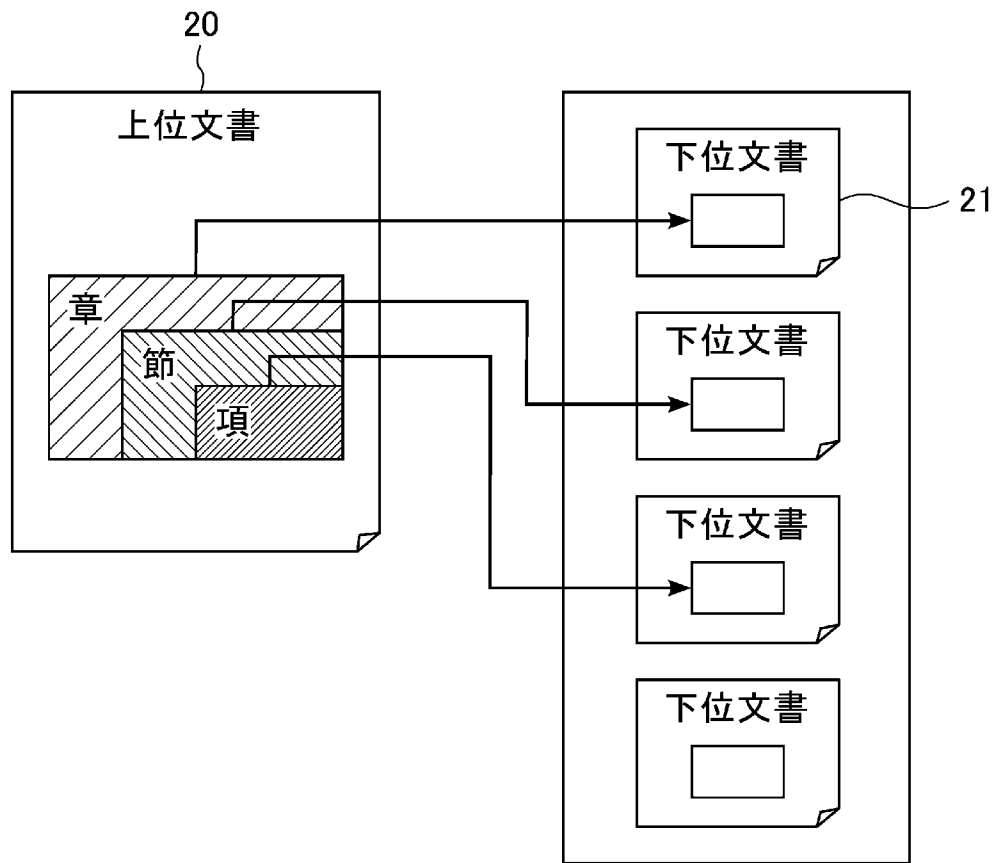
[図3]



[図4]



[図5]



[図6]

	下位文書 21_1	下位文書 21_2	下位文書 21_3	...	下位文書 21_n
上位文書20_1 の章1	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_1 の章2	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_1 の章3	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
...	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_2 の章1	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_2 の章2	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
...	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_m の章1	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
...	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1

1…対応関係有り

0…対応関係無し

[図7]

	下位文書 21_1	下位文書 21_2	下位文書 21_3	...	下位文書 21_n
上位文書20_1 章1の節1	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_1 章1の節2	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_1 章1の節3	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
...	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_1 章2の節1	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_1 章2の節2	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
...	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_2 章1の節1	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
...	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_m 章...の節...	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1

1...対応関係有り

0...対応関係無し

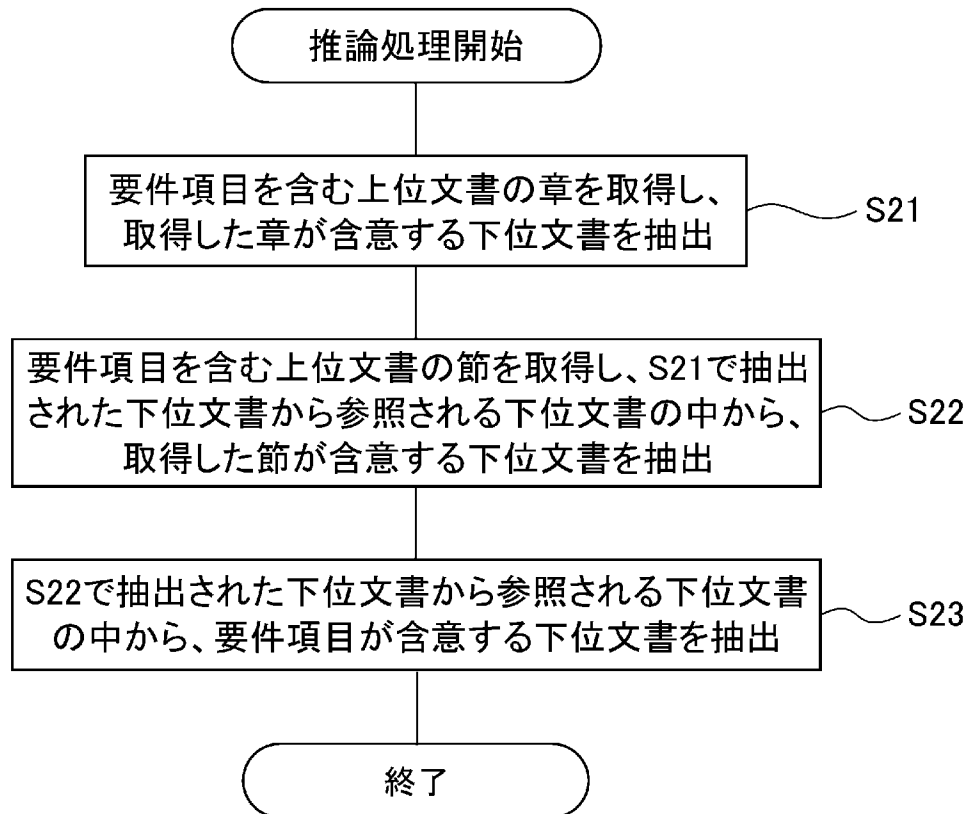
[図8]

	下位文書 21_1	下位文書 21_2	下位文書 21_3	...	下位文書 21_n
上位文書20_1 章1節1の項1	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_1 章1節1の項2	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_1 章1節1の項3	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
...	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_1 章1節2の項1	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_1 章1節2の項2	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
...	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_1 章1節3の項1	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
...	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_1 章2節1の項1	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
...	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_2 章1節1の項1	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
...	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1
上位文書20_m 章...節...の項...	0 or 1	0 or 1	0 or 1	0 or 1	0 or 1

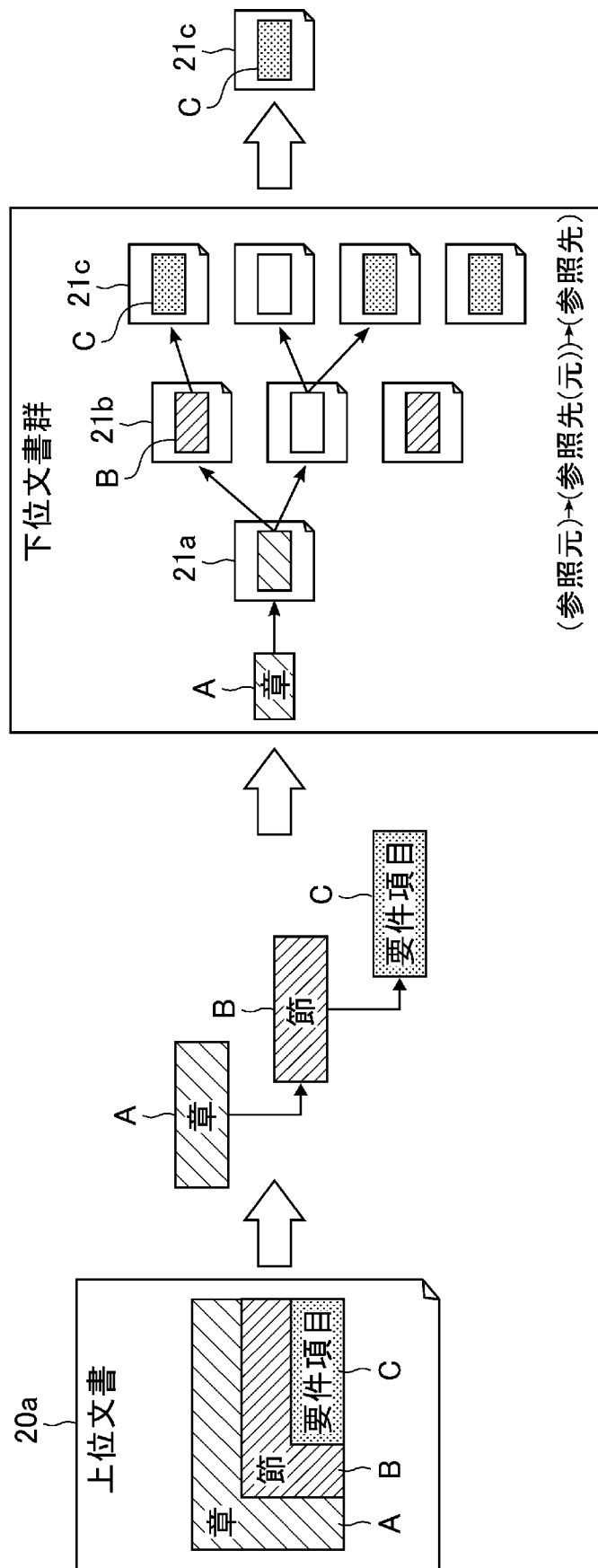
1...対応関係有り

0...対応関係無し

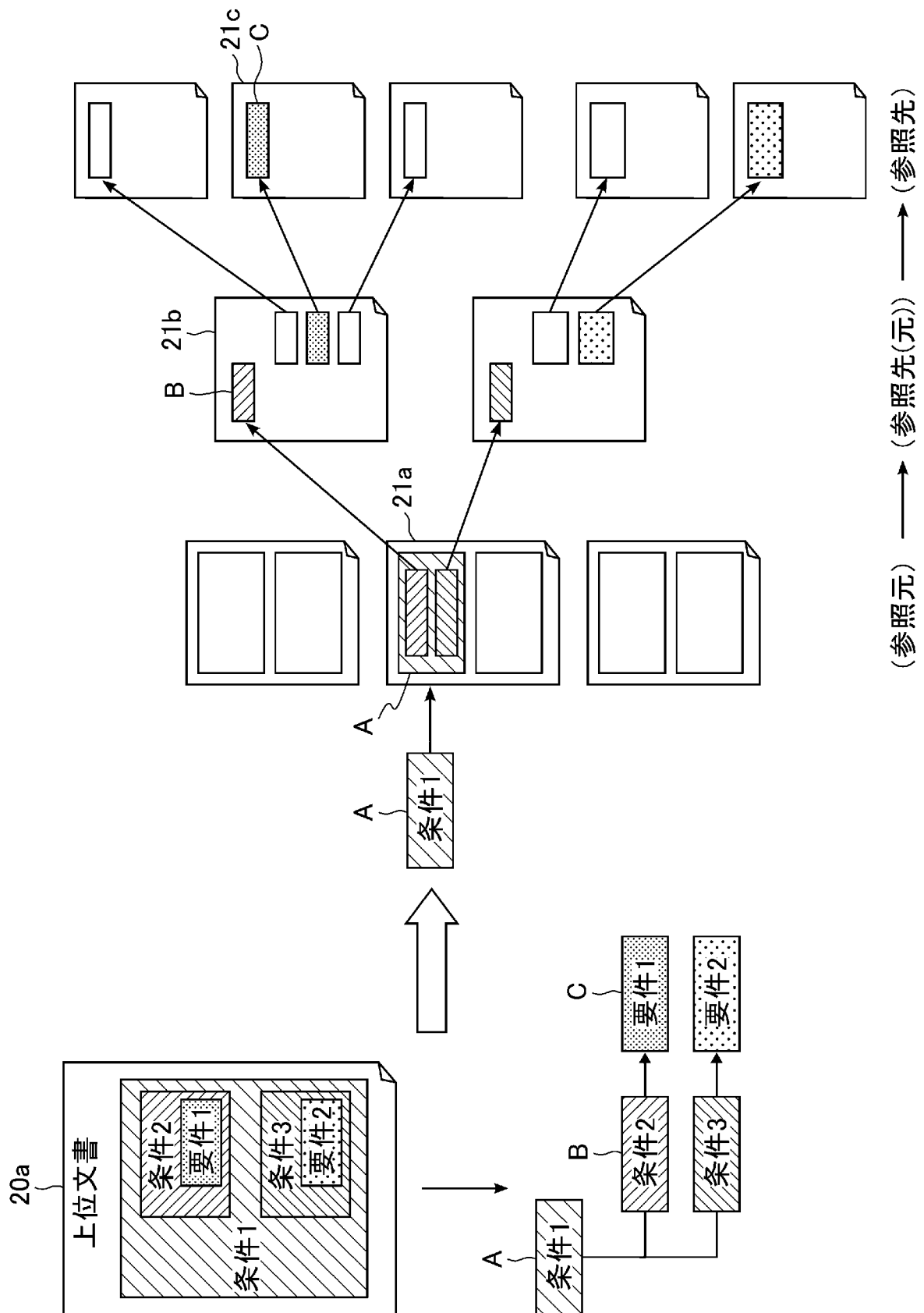
[図9]



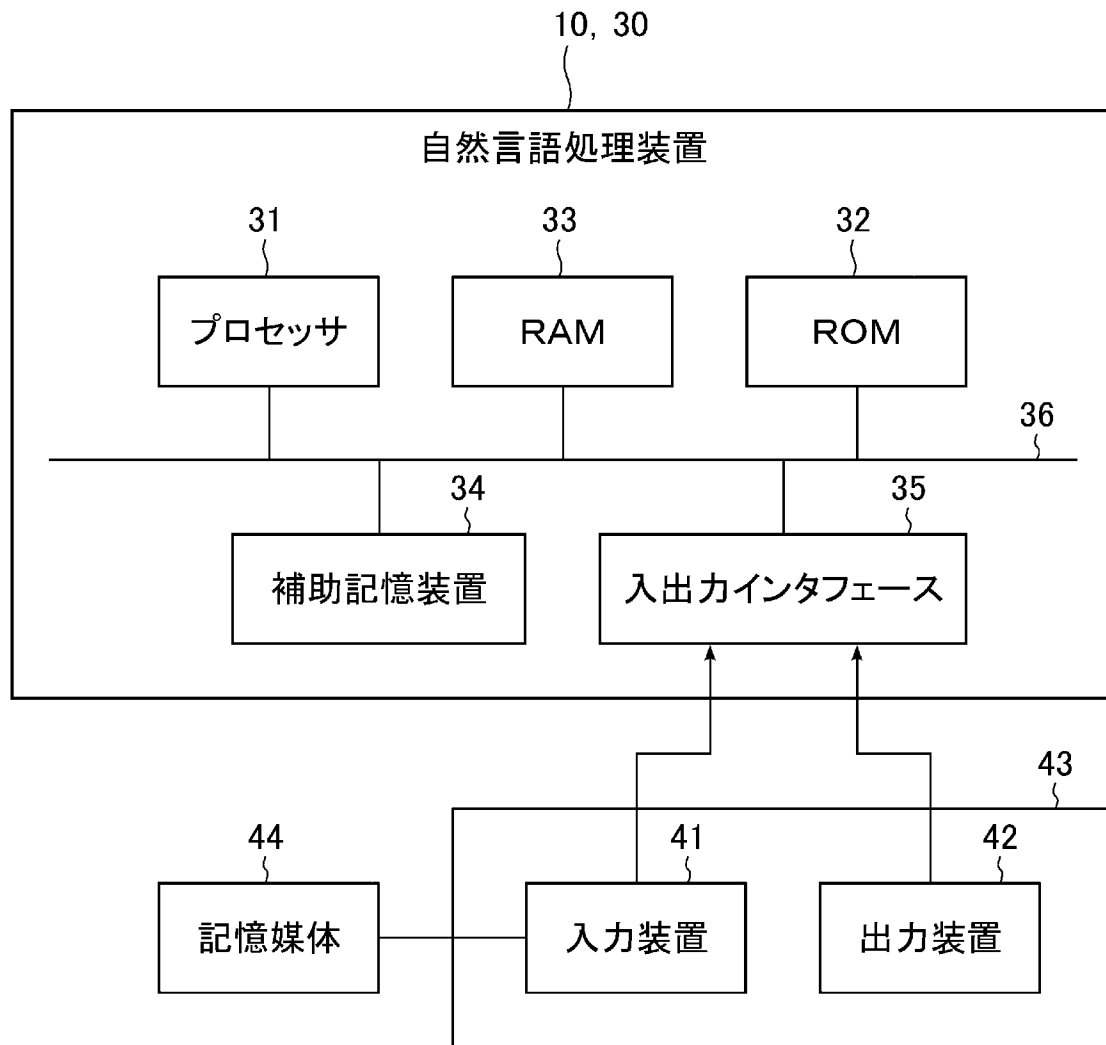
[図10]



[図11]



[図12]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2023/046002

A. CLASSIFICATION OF SUBJECT MATTER**G06F 16/33**(2019.01)i

FI: G06F16/33

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F16/33

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2024

Registered utility model specifications of Japan 1996-2024

Published registered utility model applications of Japan 1994-2024

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	中西 弘毅 ほか, 情報システムに関する内部統制文書への自然言語処理技術 BERT 適用性評価について, 電子情報通信学会技術研究報告 [online], 09 March 2023, vol. 122, no. 442, pp. 111-116, [retrieved on 11 May 2023], Internet<URL:https://ken.ieice.org/ken/user/index.php?cmd=download&p=uF0h&t=IEICE-ICM&l=7e293da2c51d4e04c14c1cc7d8df42f69580702bb604e3b3a9e83446612b997a&lang=>, (NAKANISHI, Kouki et al., Application Evaluation of Natural Language Processing Technology BERT to Internal Control Documents Related to Information Systems, IEICE Technical Report, The Institute of Electronics, Information and Communication Engineers) Introduction, sections 1, 2, 5, fig. 1-3	1-4
A	JP 2015-219799 A (HITACHI HIGH-TECHNOLOGIES CORPORATION) 07 December 2015 (2015-12-07) paragraphs [0001], [0032]-[0037], fig. 5	1-4
A	WO 2017/149711 A1 (HITACHI, LTD.) 08 September 2017 (2017-09-08) paragraphs [0004]-[0005], [0058]-[0062], fig. 9	1-4



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“D” document cited by the applicant in the international application

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

21 February 2024

Date of mailing of the international search report

05 March 2024

Name and mailing address of the ISA/JP

Japan Patent Office (ISA/JP)
3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915
Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/JP2023/046002

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP	2015-219799	A	07 December 2015	(Family: none)	
WO	2017/149711	A1	08 September 2017	(Family: none)	

A. 発明の属する分野の分類（国際特許分類（IPC）） G06F 16/33(2019.01)i FI: G06F16/33		
B. 調査を行った分野 調査を行った最小限資料（国際特許分類（IPC）） G06F16/33 最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1922 - 1996年 日本国公開実用新案公報 1971 - 2024年 日本国実用新案登録公報 1996 - 2024年 日本国登録実用新案公報 1994 - 2024年 国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）		
C. 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	中西 弘毅 ほか、情報システムに関する内部統制文書への自然言語処理技術BERT適用性評価について、電子情報通信学会技術研究報告 [online], 2023.03.09, 第122巻, 第442号, pp.111~116, [検索日 2023.05.11], Internet<URL:https://ken.ieice.org/ken/user/index.php?cmd=download&p=uF0h&t=IEICE-ICM&l=7e293da2c51d4e04c14c1cc7d8df42f69580702bb604e3b3a9e83446612b997a&lang=> (NAKANISHI, Kouki, Application Evaluation of Natural Language Processing Technology BERT to Internal Control Documents Related to Information Systems, IEICE Technical Report, The Institute of Electronics, Information and Communication Engineers.) あらまし, 第1節, 第2節, 第5節, 図1-3	1-4
A	JP 2015-219799 A (株式会社日立ハイテクノロジーズ) 07.12.2015 (2015 - 12 - 07) 段落[0001], [0032]-[0037], 図5	1-4
A	WO 2017/149711 A1 (株式会社日立製作所) 08.09.2017 (2017 - 09 - 08) 段落[0004]-[0005], [0058]-[0062], 図9	1-4
☐ C欄の続きにも文献が列挙されている。 ☒ パテントファミリーに関する別紙を参照。		
* 引用文献のカテゴリー “A” 特に関連のある文献ではなく、一般的技術水準を示すもの “D” 国際出願で出願人が先行技術文献として記載した文献 “E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの “L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す） “O” 口頭による開示、使用、展示等に言及する文献 “P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献 “T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの “X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの “Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの “&” 同一パテントファミリー文献		
国際調査を完了した日 21.02.2024	国際調査報告の発送日 05.03.2024	
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号	権限のある職員（特許庁審査官） 松尾 真人 5N 8384 電話番号 03-3581-1101 内線 3586	

国際調査報告
パテントファミリーに関する情報

国際出願番号

PCT/JP2023/046002

引用文献			公表日	パテントファミリー文献	公表日
JP	2015-219799	A	07.12.2015	(ファミリーなし)	
WO	2017/149711	A1	08.09.2017	(ファミリーなし)	