

(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
2024年9月19日(19.09.2024)



(10) 国際公開番号

WO 2024/189883 A1

(51) 国際特許分類:
G06N 5/022 (2023.01) G06N 20/00 (2019.01)

(21) 国際出願番号: PCT/JP2023/010286

(22) 国際出願日: 2023年3月16日(16.03.2023)

(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: 日本電気株式会社 (NEC CORPORATION) [JP/JP]; 〒1088001 東京都港区芝五丁目7番1号 Tokyo (JP).

(72) 発明者: 川田 拓也 (KAWADA, Takuya); 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP). 秋元 康佑 (AKIMOTO, Kosuke); 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP). 定政 邦彦 (SADAMASA, Kunihiko); 〒1088001 東京都港区芝五丁目7番1号 日本電気株式会社内 Tokyo (JP).

(74) 代理人: 弁理士法人 H A R A K E N Z O W O R L D P A T E N T & T R A D E M A R K (HARAKENZO WORLD PATENT & TRADEMARK); 〒5300041 大阪府

大阪市北区天神橋2丁目北2番6号 大和南森町ビル Osaka (JP).

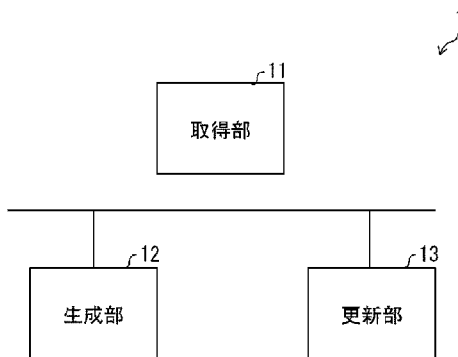
(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, 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.

(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).

(54) Title: INFORMATION PROCESSING DEVICE, INFORMATION PROCESSING METHOD, AND PROGRAM

(54) 発明の名称: 情報処理装置、情報処理方法、およびプログラム

図 1



11 Acquisition unit
12 Generation unit
13 Update unit

(57) Abstract: In order to solve the problem of improving the accuracy of relational knowledge data referenced as case examples by a language model, an information processing device (1) comprises: an acquisition unit (11) that acquires a first relational knowledge data set consisting of first relational knowledge data indicating a plurality of pieces of knowledge having a specific relationship, and also acquires information specifying knowledge context; a generation unit (12) that uses a language model to generate a second relational knowledge data set consisting of second relational knowledge data according to the context by referring to the first relational knowledge data set as case examples; and an update unit (13) that updates the first relational knowledge data set using second relational knowledge data having a specific relationship in the second relational knowledge data set.

添付公開書類：

一 国際調査報告（条約第21条(3)）

(57) 要約：言語モデルによって事例として参照される関係知識データの精度を向上させる、との課題を解決するために、情報処理装置（1）は、特定の関係を有する複数の知識を示す第1関係知識データによって構成される第1関係知識データセット、および、知識のコンテキストを指定する情報を取得する取得部（11）と、言語モデルを用いて、第1関係知識データセットを事例として参照することにより、コンテキストに応じた第2関係知識データによって構成される第2関係知識データセットを生成する生成部（12）と、第2関係知識データセットのうち、特定の関係を有する第2関係知識データを用いて、第1関係知識データセットを更新する更新部（13）と、を含む。

明 細 書

発明の名称：情報処理装置、情報処理方法、およびプログラム

技術分野

[0001] 本発明は、関係知識を構築する技術に関する。

背景技術

[0002] 互いに関係を有する複数の知識は、関係知識と呼称される。例えば、或る業界における関係知識は、当該業界の状況を把握したり、将来予測を行ったりするための有用な情報となり得る。非特許文献1には、言語処理タスクの事例を与えることで目的の言語処理タスクを実行可能な言語モデルを用いて、少量の関係知識データを事例として参照して新たな関係知識データを生成する技術が記載されている。

先行技術文献

非特許文献

[0003] 非特許文献1：Peter West et. al., “Symbolic Knowledge Distillation: from General Language Models to Commonsense Models” arXiv: 2110.07178 v2 [cs.CL] 28 Nov 2022

発明の概要

発明が解決しようとする課題

[0004] しかしながら、非特許文献1に記載された技術においては、生成される関係知識データは、或るコンテキストにおいては適切であっても、異なるコンテキストにおいては適切でない可能性があるという問題があった。また、生成される関係知識データの精度は、参照する事例の精度に依存する可能性があるという問題があった。このような問題を解決するためには、言語モデルによって事例として参照される関係知識データの精度を向上させることが重要である。

[0005] 本発明の一態様は、上記の問題に鑑みてなされたものであり、その目的の一例は、言語モデルによって事例として参照される関係知識データの精度を

向上させることである。

課題を解決するための手段

- [0006] 本発明の一側面に係る情報処理装置は、特定の関係を有する複数の知識を示す第1関係知識データによって構成される第1関係知識データセット、および、知識のコンテキストを指定する情報を取得する取得手段と、言語モデルを用いて、前記第1関係知識データセットを事例として参照することにより、前記コンテキストに応じた第2関係知識データによって構成される第2関係知識データセットを生成する生成手段と、前記第2関係知識データセットのうち、前記特定の関係を有する第2関係知識データを用いて、前記第1関係知識データセットを更新する更新手段と、を含む。
- [0007] 本発明の一側面に係る情報処理方法は、1以上のプロセッサが実行する情報処理方法であって、特定の関係を有する複数の知識を示す第1関係知識データによって構成される第1関係知識データセット、および、知識のコンテキストを指定する情報を取得することと、言語モデルを用いて、前記第1関係知識データセットを事例として参照することにより、前記コンテキストに応じた第2関係知識データによって構成される第2関係知識データセットを生成することと、前記第2関係知識データセットのうち、前記特定の関係を有する第2関係知識データを用いて、前記第1関係知識データセットを更新することと、を含む。
- [0008] 本発明の一側面に係るプログラムは、コンピュータを、特定の関係を有する複数の知識を示す第1関係知識データによって構成される第1関係知識データセット、および、知識のコンテキストを指定する情報を取得する取得手段と、言語モデルを用いて、前記第1関係知識データセットを事例として参照することにより、前記コンテキストに応じた第2関係知識データによって構成される第2関係知識データセットを生成する生成手段と、前記第2関係知識データセットのうち、前記特定の関係を有する第2関係知識データを用いて、前記第1関係知識データセットを更新する更新手段と、として機能させる。

発明の効果

[0009] 本発明の一態様によれば、言語モデルによって事例として参照される関係知識データの精度を向上させることができる。

図面の簡単な説明

[0010] [図1]本発明の例示的实施形態1に係る情報処理装置の構成を示すブロック図である。

[図2]本発明の例示的实施形態1に係る情報処理方法の流れを示すフロー図である。

[図3]本発明の例示的实施形態2に係る情報処理システムの構成を示すブロック図である。

[図4]本発明の例示的实施形態2に係る情報処理方法の流れを示すフロー図である。

[図5]図4に示す情報処理方法の各ステップにおいて生成される情報の遷移を示す模式図である。

[図6]図5に続く模式図である。

[図7]第1関係知識データセットの具体例を示す図である。

[図8]コンテキストリストの具体例を示す図である。

[図9]第1プロンプトの具体例を説明する図である。

[図10]第2前件データリストの具体例を示す図である。

[図11]第2プロンプトの具体例を説明する図である。

[図12]第2関係知識データセットの具体例を示す図である。

[図13]第3関係知識データセットの具体例を示す図である。

[図14]第3プロンプトの具体例を説明する図である。

[図15]本発明の各例示的实施形態に係る情報処理装置のハードウェア構成例を示すブロック図である。

発明を実施するための形態

[0011] [例示的实施形態1]

本発明の第1の例示的实施形態について、図面を参照して詳細に説明する

。本例示的实施形態は、後述する例示的实施形態の基本となる形態である。

[0012] (情報処理装置 1 の構成)

本例示的实施形態に係る情報処理装置 1 の構成について、図 1 を参照して説明する。図 1 は、情報処理装置 1 の構成を示すブロック図である。図 1 に示すように、情報処理装置 1 は、取得部 1 1 と、生成部 1 2 と、更新部 1 3 と、を含む。

[0013] 取得部 1 1 は、特定の関係を有する複数の知識を示す第 1 関係知識データによって構成される第 1 関係知識データセット、および、知識のコンテキストを指定する情報を取得する。生成部 1 2 は、言語モデルを用いて、第 1 関係知識データセットを事例として参照することにより、コンテキストに応じた第 2 関係知識データによって構成される第 2 関係知識データセットを生成する。更新部 1 3 は、第 2 関係知識データセットのうち、特定の関係を有する第 2 関係知識データを用いて、第 1 関係知識データセットを更新する。

[0014] (プログラムによる実現例)

上述した情報処理装置 1 を、1 以上のプロセッサおよびメモリを備えるコンピュータによって構成する場合、本例示的实施形態は、当該メモリに記憶されたプログラムによって実現することができる。当該プログラムは、コンピュータを、特定の関係を有する複数の知識を示す第 1 関係知識データによって構成される第 1 関係知識データセット、および、知識のコンテキストを指定する情報を取得する取得部 1 1 と、言語モデルを用いて、第 1 関係知識データセットを事例として参照することにより、コンテキストに応じた第 2 関係知識データによって構成される第 2 関係知識データセットを生成する生成部 1 2 と、第 2 関係知識データセットのうち、特定の関係を有する第 2 関係知識データを用いて、第 1 関係知識データセットを更新する更新部 1 3 と、として機能させる。

[0015] (情報処理方法 S 1 の流れ)

以上の様に構成された情報処理装置 1 は、本例示的实施形態に係る情報処理方法 S 1 を実行する。上述した情報処理装置 1 を、1 以上のプロセッサお

よびメモリを備えるコンピュータによって構成する場合、情報処理方法 S 1 は、例えば、当該 1 以上のプロセッサが上述したプログラムを読み込んで実行することにより実現される。

[0016] 情報処理方法 S 1 の流れについて、図 2 を参照して説明する。図 2 は、情報処理方法 S 1 の流れを示すフロー図である。情報処理方法 S 1 は、ステップ S 1 1、S 1 2、S 1 3 を含む。ステップ S 1 1 において、取得部 1 1 は、特定の関係を有する複数の知識を示す第 1 関係知識データによって構成される第 1 関係知識データセット、および、知識のコンテキストを指定する情報を取得する。ステップ S 1 2 において、生成部 1 2 は、言語モデルを用いて、第 1 関係知識データセットを事例として参照することにより、コンテキストに応じた第 2 関係知識データによって構成される第 2 関係知識データセットを生成する。ステップ S 1 3 において、更新部 1 3 は、第 2 関係知識データセットのうち、特定の関係を有する第 2 関係知識データを用いて、第 1 関係知識データセットを更新する。

[0017] (本例示的实施形態の効果)

以上のように、本例示的实施形態においては、特定の関係を有する複数の知識を示す第 1 関係知識データによって構成される第 1 関係知識データセット、および、知識のコンテキストを指定する情報を取得し、言語モデルを用いて、第 1 関係知識データセットを事例として参照することにより、コンテキストに応じた第 2 関係知識データによって構成される第 2 関係知識データセットを生成し、第 2 関係知識データセットのうち、特定の関係を有する第 2 関係知識データを用いて、第 1 関係知識データセットを更新する、との構成が採用されている。このため、本例示的实施形態によれば、言語モデルによって事例として参照される関係知識データの精度を向上させることができる、という効果が得られる。

[0018] [例示的实施形態 2]

本発明の第 3 の例示的实施形態に係る情報処理装置 1 A を含む情報処理システム 1 0 0 について、図面を参照して詳細に説明する。なお、例示的实施

形態 1 にて説明した構成要素と同じ機能を有する構成要素については、同じ符号を付し、その説明を適宜省略する。

[0019] <情報処理システム 100 の概要>

情報処理システム 100 は、言語モデル 3 A を用いて、指定したコンテキストにおける特定の関係を有する関係知識データを生成するシステムである。ここで、関係知識データとは、何等かの関係を有する複数の知識によって構成される。特定の関係を有する関係知識データは、特定の関係を有する複数の知識によって構成される。また、知識は、自然言語文によって表される情報である。また、コンテキストとは、複数の知識の間で特定の関係が成立する範囲または条件を規定する情報である。コンテキストの一例としては、産業の種別、国、人、時代、等が挙げられるが、これらに限られない。本例示的实施形態においては、「特定の関係」として、「因果関係」、「時間的關係」等のように、前後関係が規定された関係を適用する。特定の関係を有する複数の知識のうち、前後関係が「前」となる知識を「前件の知識」と呼称する。また、前後関係が「後」となる知識を「後件の知識」と呼称する。換言すると、前後関係が規定された関係とは、前件の知識および後件の知識を入れ替えると元の関係が成り立たない関係である。例えば、「特定の関係」として「因果関係」を適用する場合、前件の知識は原因を示し、後件の知識は結果を示す。

[0020] <情報処理システム 100 の構成>

図 3 は、情報処理システム 100 の構成を示すブロック図である。図 3 に示すように、情報処理システム 100 は、情報処理装置 1 A と、関係知識データベース 2 A と、言語モデル 3 A とを含む。情報処理装置 1 A は、関係知識データベース 2 A および言語モデル 3 A のそれぞれと通信可能に接続される。

[0021] (関係知識データベース 2 A)

関係知識データベース 2 A は、特定の関係を有する関係知識データを記憶する装置である。本例示的实施形態における関係知識データは、前件データ

、後件データ、およびコンテキストデータを含む。前件データは、前件の知識を示す自然言語文である。後件データは、後件の知識を示す自然言語文である。コンテキストデータは、前件および後件の知識に関連するコンテキストを示す自然言語文である。

[0022] (言語モデル 3 A)

言語モデル 3 A は、言語処理タスクの事例を参照することにより当該言語処理タスクを実行するように学習された汎用的な大規模言語モデルである。以降では、事例をサンプルとも記載する。例えば、言語モデル 3 A は、プロンプトを入力として当該プロンプトに応じた情報を生成するモデルであってもよい。プロンプトおよび生成される情報は、それぞれ自然言語文である。プロンプトは、インストラクション、サンプル、およびクエリを含む。インストラクションは、言語モデル 3 A に生成させたい情報を、当該言語モデル 3 A に教示する内容を示す。言語モデル 3 A は、インストラクションに基づきサンプルを事例として参照することにより、クエリに応じた情報を生成する。また、言語モデル 3 A は、生成した情報と共に、当該情報の適切さを示すスコアを出力するものであってもよい。言語モデル 3 A の一例としては、G T P 3 (Generative Pre-trained Transformer 3) が挙げられるが、これに限られない。

[0023] (情報処理装置 1 A)

図 3 に示すように、情報処理装置 1 A は、制御部 1 1 0 と、記憶部 1 2 0 と、を含む。制御部 1 1 0 は、情報処理装置 1 A の各部を統括して制御する。記憶部 1 2 0 は、制御部 1 1 0 が使用する各種データを記憶する。制御部 1 1 0 は、取得部 1 1 A と、生成部 1 2 A と、判定部 1 6 A と、更新部 1 3 A と、を含む。

[0024] 取得部 1 1 A は、例示的实施形態 1 における取得部 1 1 と同様に、第 1 関係知識データセットおよびコンテキストを指定する情報を取得する。図 3 に示すコンテキストリストは、コンテキストを指定する情報の一例であり、1 または複数のコンテキストを含んでいる。また、本例示的实施形態では、第

1 関係知識データセットを構成する第1 関係知識データは、前件の知識を示す第1 前件データ、および、前件の知識に対して特定の関係を有する後件の知識を示す第1 後件データを含む。

[0025] 生成部1 2 Aは、言語モデル3 Aを用いて、第1 関係知識データセットを事例として参照することにより、コンテクストリストに応じた第2 関係知識データセットを生成する。また、生成部1 2 Aは、更新部1 3 Aにより更新された第1 関係知識データセットを事例として参照することにより、新たな第2 関係知識データセットを生成する。

[0026] 例えば、生成部1 2 Aは、第1 生成部1 4 Aおよび第2 生成部1 5 Aを含む。第1 生成部1 4 Aは、第1 前件データを事例として参照することにより、コンテクストに応じた第2 前件データを生成する。第2 生成部1 5 Aは、第1 関係知識データを事例として参照することにより、当該第2 前件データに応じた第2 後件データを生成する。これにより、生成部1 2 Aは、第2 前件データおよび第2 後件データを含む第2 関係知識データを生成する。

[0027] 判定部1 6 Aは、言語モデル3 Aを用いて、第3 関係知識データによって構成される第3 関係知識データセットを事例として参照することにより、第2 関係知識データが当該異なる関係を有するか否かを判定する。第3 関係知識データは、特定の関係とは異なる関係を有する複数の知識を示す。

[0028] 更新部1 3 Aは、例示的实施形態1 における更新部1 3 と同様に構成されることに加えて、第1 関係知識データセットを更新するために用いる「特定の関係を有する第2 関係知識データ」を次のようにして特定する。例えば、更新部1 3 Aは、第2 関係知識データセットのうち、言語モデル3 Aが算出したスコアに基づく条件を満たす第2 関係知識データを、特定の関係を有する第2 関係知識データとして用いて、第1 関係知識データセットを更新する。また、更新部1 3 Aは、第2 関係知識データセットのうち、ユーザによって選択された第2 関係知識データを、特定の関係を有する第2 関係知識データとして用いて、第1 関係知識データセットを更新する。

[0029] また、更新部1 3 Aは、第2 関係知識データセットのうち、上述した「異

なる関係」を有していないと判定された第2関係知識データを、特定の関係を有する第2関係知識データとして用いて、第1関係知識データセットを更新する。なお、更新部13Aは、第1関係知識データセットの更新に用いる「特定の関係を有する第2関係知識データ」を特定するために、スコアに基づく条件、ユーザによる選択、および判定部16Aによる判定の何れか1つを用いてもよいし、一部または全部の組み合わせを用いてもよい。

[0030] また、更新部13Aは、第2関係知識データセットのうち、上述した「異なる関係」を有していると判定された第2関係知識データを用いて、第3関係知識データセットを更新する。なお、更新部13Aは、第3関係知識データセットの更新に用いる「異なる関係を有する第2関係知識データ」を特定するために、判定部16Aによる判定に替えて、または加えて、スコアに基づく条件、ユーザによる選択、またはこれらの組み合わせ等を用いてもよい。

[0031] 記憶部120は、第1インストラクションデータと、第2インストラクションデータと、第3インストラクションデータと、第2関係知識データセットと、第3関係知識データセットとを記憶する。記憶部120が記憶する情報のうち、第1インストラクションデータ、第2インストラクションデータ、第3インストラクションデータおよび第3関係知識データセットは、後述する情報処理方法S1Aの実行前に記憶されており、情報処理方法S1Aの実行により更新される。第2関係知識データセットは、当該情報処理方法S1Aの実行前に記憶されている必要はなく、情報処理方法S1Aの実行により生成される。また、情報処理装置1Aには、第1関係知識データセットおよびコンテクストリストが入力される。

[0032] <情報処理方法S1Aの流れ>

以上のように構成された情報処理装置1Aは、本例示的实施形態に係る情報処理方法S1Aを実行する。情報処理方法S1Aについて、図4～図6を参照して説明する。図4は、情報処理方法S1Aの流れを示すフロー図である。図5は、情報処理方法S1Aの各ステップにおいて生成される情報の遷

移を示す模式図である。図6は、図5に続く模式図である。

[0033] (ステップS101)

ステップS101において、取得部11Aは、第1関係知識データセットおよびコンテキストリストを取得する。

[0034] (第1関係知識データセット)

第1関係知識データセットは、言語モデル3Aにサンプルとして入力される情報を含む。情報処理方法S1Aは繰り返し実行され得るが、初回実行時に取得される第1関係知識データセットは、人が作成したものであってもよいし、他の装置が作成したものであってもよい。また、2回目以降に取得される第1関係知識データセットは、前回の情報処理方法S1Aの実行により更新されたものである。

[0035] 第1関係知識データセットは、1または複数の第1関係知識データによって構成される。第1関係知識データは、特定の関係を有する。本例示的实施形態では、第1関係知識データは、第1前件データ、第1後件データ、および第1コンテキストデータを含む。

[0036] 図7は、第1関係知識データセットの具体例を示す図である。図7に示すように、第1関係知識データセットは、第1関係知識データD11、D12、D13、…を含む。これらを特に区別して説明する必要がない場合には、それぞれを第1関係知識データD1とも記載する。第1関係知識データD1は、「特定の関係」の一例である「因果関係」を有する。第1関係知識データD1は、第1項目「INDUSTRY」、第2項目「CAUSE」、および第3項目「EFFECT」を含む。第1関係知識データD1において項目名「INDUSTRY」に続く自然言語文は、第1コンテキストデータの一例であり、産業の種別を示す。項目名「CAUSE」に続く自然言語文は、第1前件データの一例であり、当該産業において原因となる前件の知識を示す。項目名「EFFECT」に続く自然言語文は、第1後件データの一例であり、当該原因に対する結果となる後件の知識を示す。

[0037] 例えば、第1関係知識データD11は、第1コンテキストデータが示すコ

ンテキスト「agriculture」において、第1前件データが示す前件の知識「There is a growing focus on functional fruits / Can clearly state the relationship between fruits and health.」と、第1後件データが示す後件の知識「Enhance and add a variety of nutrients to fruits.」とが、因果関係を有することを示している。

[0038] なお、図7では、第1関係知識データD12、D13については、第1コンテキストデータ、第1前件データ、および第1後件データとして「×××××」等といった意味を持たない文字列を示しているが、これらの文字列は、自然言語文を模式的に表すものである。以降の図面においても、このような意味を持たない文字列は、自然言語文を模式的に表すものとする。

[0039] (コンテキストリスト)

コンテキストリストは、コンテキストを指定する情報の一例であり、1または複数のコンテキストを指定する情報を含む。以降、コンテキストリストが示す各コンテキストを「指定されたコンテキスト」とも記載し、指定されたコンテキストを示す情報を「指定コンテキストデータ」とも記載する。指定コンテキストデータは、言語モデル3Aにクエリとして入力される。図8は、コンテキストリストの具体例を示す図である。図8の例では、コンテキストリストは、「agriculture」、「fishing」、「forestry」、…等といったように、自然言語文で表された複数の指定コンテキストデータを含んでいる。

[0040] (ステップS102～S103：第2前件データの生成)

再び図4～図5を参照して、ステップS102～S103について説明する。第1生成部14Aは、コンテキストリストにより指定された各コンテキストについてステップS102～S103を実行することにより、指定された各コンテキストに応じた第2前件データによって構成される第2前件データリストを生成する。具体的には、まず、ステップS102において、第1生成部14Aは、言語モデル3Aに入力するための第1プロンプトを生成する。

[0041] (第1プロンプト)

第1プロンプトは、指定されたコンテキストに応じた第2前件データを生成するために言語モデル3Aに入力する情報である。図5を参照して、第1プロンプトの構成について説明する。図5に示すように、ステップS102において生成される第1プロンプトは、「インストラクション」として第1インストラクションデータと、「サンプル」として第1前件データおよび第1コンテキストデータと、「クエリ」として指定コンテキストデータと、を含む。

[0042] 例えば、第1生成部14Aは、第1関係知識データセットから一部または全部の第1関係知識データを抽出し、抽出した各第1関係知識データに含まれる第1コンテキストデータおよび第1前件データをサンプルとして第1プロンプトに含める。また、例えば、第1生成部14Aは、指定コンテキストデータを「クエリ」として第1プロンプトに含める。また、例えば、第1生成部14Aは、記憶部120に記憶された第1インストラクションデータを、「インストラクション」として第1プロンプトに含める。第1インストラクションデータは、第1プロンプトにより生成したい情報を言語モデル3Aに教示する内容を示す自然言語文であり、テンプレートとして事前に記憶部120に記憶されている。

[0043] 図9は、第1プロンプトの具体例を説明する図である。図9の例では、第1インストラクションデータは、当該第1プロンプトにより生成したい情報が第1項目「INDUSTRY」および第2項目「STEEP」を含むこと、並びに、各項目の定義を説明する自然言語文を含む。なお、項目名「STEEP」は、第1前件データに対応する項目名であればよく、上述した例に限られない。

[0044] また、図9の例では、第1プロンプトに含まれるサンプルは、サンプルSMP11を含む。サンプルSMP11は、図7に示した第1関係知識データD12における第1コンテキストデータを第1項目「INDUSTRY」として含み、第1前件データを第2項目「STEEP」として含む。当該サン

プルSMP 11を生成するための第1関係知識データD12は、第1生成部14Aによって第1関係知識データセットから抽出されたものであってよい。当該抽出処理は、ランダムに行われてもよいし、所定の条件を満たすものが抽出されてもよい。例えば、所定の条件とは、スコア、生成日時、過去の抽出履歴、等に基づく条件であってもよい。

[0045] なお、ステップS102における第1プロンプトの生成処理は、コンテキストリストにより指定された各コンテキストについて個別に実行されてもよいし、一部または全部の複数のコンテキストについて共通して1回実行されてもよい。個別に実行される場合、指定された各コンテキストについて生成される第1プロンプトに含まれるサンプルは、互いに異なるものとなり得る。これにより、サンプルの内容的な偏りが低減され、より多様な第2前件データを生成することができる。なお、図9の例では、第1プロンプトに含まれる1つのサンプルSMP 11を示しているが、第1プロンプトに含まれるサンプルの個数は、複数であってもよい。

[0046] また、図9の例では、第1プロンプトに含まれるクエリは、第1項目「INDUSTRY」として指定コンテキストデータを含むが、第2項目「STEEP」として項目名のみを含み当該項目名に続く自然言語文を含まない。このようなクエリは、指定コンテキストデータが示すコンテキスト「fishing」に対応する「STEEP」を示す自然言語文の生成要求を示している。

[0047] 次に、ステップS103において、第1生成部14Aは、生成した第1プロンプトを言語モデル3Aに入力することにより、指定されたコンテキストに応じた第2前件データを生成する。ステップS102～S103の処理を、コンテキストリストにより指定された各コンテキストについて実行することにより、図5に示すように第2前件データリストが生成される。

[0048] (第2前件データリスト)

第2前件データリストは、指定された各コンテキストについて生成された1または複数の第2前件データを含む。図10は、第2前件データリストの具体例を示す図である。図10に示すように、第2前件データリストは、第

2 前件データと、当該第 2 前件データが示す知識のコンテキストを示す第 2 コンテキストデータと、を含んでいる。第 2 コンテキストデータとしては、当該第 2 前件データを生成するために第 1 プロンプトにクエリとして含めた指定コンテキストデータが適用される。図 10 の 1 行目は、例えば、言語モデル 3 A に対して、図 9 に示した第 1 プロンプトが入力されることにより生成される。具体的には、言語モデル 3 A は、図 9 に示す第 1 インストラクションデータに基づきサンプルを事例として、クエリに含まれる第 1 項目「fishing」に応じた第 2 項目「The use of robots for fishing purposes is becoming more widespread. Robots can be used for a variety of tasks.」を生成する。第 2 項目データとして生成された自然言語文は、第 2 前件データの一例となる。つまり、第 1 前件データをサンプルとすることにより、指定されたコンテキスト「fishing」において何らかの原因となり得る「前件の知識」が生成される。なお、第 1 生成部 1 4 A は、1 つのコンテキストについてステップ S 1 0 3 を複数回繰り返してもよい。これにより、1 つのコンテキストについて複数の第 2 前件データが生成されてもよい。

[0049] (ステップ S 1 0 4 ~ S 1 0 5 : 第 2 関係知識データセットの生成)

再び図 4 ~ 図 5 を参照して、ステップ S 1 0 4 ~ S 1 0 5 について説明する。第 2 生成部 1 5 A は、第 2 前件データリストに含まれる各第 2 前件データについてステップ S 1 0 4 ~ S 1 0 5 を実行することにより、各第 2 前件データを含む第 2 関係知識データによって構成される第 2 関係知識データセットを生成する。具体的には、まず、ステップ S 1 0 4 において、第 1 生成部 1 4 A は、言語モデル 3 A に入力するための第 2 プロンプトを生成する。

[0050] (第 2 プロンプト)

第 2 プロンプトは、指定された第 2 前件データに対応する第 2 後件データを生成するために言語モデル 3 A に入力する情報である。図 5 を参照して、第 2 プロンプトの構成について説明する。図 5 に示すように、ステップ S 1 0 4 において生成される第 2 プロンプトは、「インストラクション」として第 2 インストラクションデータと、「サンプル」として第 1 関係知識データ

と、「クエリ」として第2コンテキストデータおよび第2前件データと、を含む。

[0051] 例えば、ステップS104において、第1生成部14Aは、第1関係知識データセットから一部または全部の第1関係知識データを抽出し、抽出した各第1関係知識データをサンプルとして第2プロンプトに含める。また、例えば、第1生成部14Aは、第2前件データリストに含まれる第2コンテキストデータおよび第2前件データを「クエリ」として第2プロンプトに含める。また、例えば、第2生成部15Aは、記憶部120に記憶された第2インストラクションデータを「インストラクション」として第2プロンプトに含める。第2インストラクションデータは、第2プロンプトにより生成したい情報を言語モデル3Aに教示する内容を示す自然言語文であり、テンプレートとして事前に記憶部120に記憶されている。

[0052] 図11は、第2プロンプトの具体例を説明する図である。図11の例では、第2インストラクションデータは、当該第2プロンプトにより生成したい情報が第1項目「INDUSTRY」、第2項目「CAUSE」、および第3項目「EFFECT」の各項目を含むこと、並びに、各項目の定義を説明する自然言語文を含む。

[0053] また、図11の例では、第2プロンプトに含まれるサンプルは、サンプルSMP21を含む。サンプルSMP21は、第1項目「INDUSTRY」、第2項目「CAUSE」および第3項目「EFFECT」として図7に示した第1関係知識データD13を含む。当該サンプルSMP21を生成するための第1関係知識データD13は、第2生成部15Aによって第1関係知識データセットから抽出されたものであってよい。当該抽出処理は、ランダムに行われてもよいし、所定の条件を満たすものが抽出されてもよい。例えば、所定の条件とは、スコア、生成日時、過去の抽出履歴、等に基づく条件であってもよい。

[0054] なお、ステップS104における第2プロンプトの生成処理は、第2前件データリストに含まれる各第2前件データについて個別に実行されてもよい

し、一部または全部の複数の第2前件データについて共通して1回実行されてもよい。個別に実行される場合、各第2前件データについて生成される第2プロンプトに含まれるサンプルは、互いに異なるものとなり得る。これにより、サンプルの内容的な偏りが低減され、より多様な第2後件データを生成することができる。なお、図11の例では、第2プロンプトに含まれる1つのサンプルSMP21を示しているが、第2プロンプトに含まれるサンプルの個数は、複数であってもよい。

[0055] また、図11の例では、第2プロンプトに含まれるクエリは、第1項目「INDUSTRY」として第2コンテキストデータを含み、第2項目「CAUSE」として第2前件データを含む。一方、当該クエリは、第3項目「EFFECT」として項目名のみを含み当該項目名に続く自然言語文を含まない。このようなクエリは、第2コンテキストデータが示すコンテキスト「fishing」および第2前件データ「The use of robots …」に対応する「EFFECT」を示す自然言語文の生成要求を示している。

[0056] 次に、ステップS105において、第2生成部15Aは、生成した第2プロンプトを言語モデル3Aに入力することにより、第2コンテキストデータおよび第2前件データに応じた第2後件データを生成する。ステップS104～S105の処理を、第2前件データリストに含まれる各第2前件データについて実行することにより、図5に示すように、指定コンテキストに応じた第2関係知識データセットが生成される。

[0057] (第2関係知識データセット)

第2関係知識データセットは、第2前件データリストに含まれる1または複数の第2前件データと、当該第2前件データに対応する第2コンテキストデータと、当該第2前件データに対応する第2後件データと、を含む。図12は、第2関係知識データセットの具体例を示す図である。図12に示すように、第2関係知識データセットは、第2関係知識データD21、D22、…を含む。これらを特に区別して説明する必要がない場合には、それぞれを第2関係知識データD2とも記載する。第2関係知識データD2は、第2コ

ンテキストデータ、第2前件データ、および第2後件データを含む。換言すると、第2関係知識データセットは、図10に示した第2前件データリストに「第2後件データ」の列を追加した構成である。第2関係知識データD2は、第1関係知識データD1をサンプルとして言語モデル3Aが生成したものであるため、第1関係知識データD1と同様に「特定の関係」を有する可能性が高いが、必ずしも「特定の関係」を有しているとは限らない。そこで、第2関係知識データが有する関係に応じた処理を行うために続くステップS106～S108が実行される。

[0058] (ステップS106～S108：第2関係知識データが有する関係を判定)

再び図4、図6を参照して、ステップS106～S108について説明する。判定部16Aは、これらのステップを実行することにより、各第2関係知識データが有する関係が「特定の関係」、「異なる関係」、および「それ以外の関係」の何れであるかを、言語モデル3Aを用いて判定する。以降では、「特定の関係」を有する複数の知識を正例とも記載し、「異なる関係」を有する複数の知識を負例とも記載する。具体的には、まず、ステップS106において、判定部16Aは、言語モデル3Aに入力するための第3プロンプトを生成する。また、第3プロンプトの生成においては、第3関係知識データセットが用いられる。

[0059] (第3関係知識データセット)

第3関係知識データセットは、言語モデル3Aにサンプルとして入力される情報を含む。情報処理方法S1Aは繰り返し実行され得るが、初回実行前に記憶部120に記憶されている第3関係知識データセットは、人が作成したものであってもよいし、他の装置が作成したものであってもよい。また、2回目以降に用いられる第3関係知識データセットは、前回の情報処理方法S1Aの実行により更新されたものである。

[0060] 第3関係知識データセットは、1または複数の第3関係知識データを示す。前述したように、第3関係知識データは、特定の関係とは異なる関係を有

する複数の知識を示す。本例示の実施形態では、第3関係知識データは、第3コンテキストデータ、第3前件データ、および第3後件データを含む。第3前件データおよび第3後件データは、第3コンテキストデータが示すコンテキストにおいて「異なる関係」を有する。例えば、「特定の関係」として「因果関係」を適用する場合、「異なる関係」は、「同義関係」、「含意関係」等であってもよいが、これらに限られない。ここで、「異なる関係」は、「同義関係」、「含意関係」のように前後関係が規定されない関係であってもよい。前後関係が規定されない関係とは、第3前件データおよび第3後件データを入れ替えても当該「異なる関係」が成り立つ関係である。ただし、「異なる関係」はこれに限らず、前後関係が規定された関係であってもよい。

[0061] 図13は、第3関係知識データセットの具体例を示す図である。図13に示すように、第3関係知識データセットは、第3関係知識データD31、D32、D33、…を含む。これらを特に区別して説明する必要がない場合には、それぞれを第3関係知識データD3とも記載する。第3関係知識データD3は、「異なる関係」の一例である「同義関係」を有する。第3関係知識データD3は、第1項目「INDUSTRY」、第2項目「SENTENCE1」、および第3項目「SENTENCE2」を含む。項目名「INDUSTRY」に続く自然言語文は、第3コンテキストデータの一例であり、産業の種別を示す。項目名「SENTENCE1」に続く自然言語文は、第3前件データの一例であり、「異なる関係」を有する複数の知識の一方を示す。項目名「SENTENCE2」に続く自然言語文は、第3後件データの一例であり、「異なる関係」を有する複数の知識の他方を示す。

[0062] 例えば、第3関係知識データD31は、第3コンテキストデータが示すコンテキスト「Nursing Care Business」において、第3前件データが示す知識「The use of robots in nursing care is becoming more widespread.」と、第3後件データが示す知識「Recently the use of robots in nursing care is very popular.」とが、同義関係を有することを示している。

[0063] (第3プロンプト)

第3プロンプトは、第2関係知識データが「異なる関係」を有する負例であるか否かを判定するために言語モデル3Aに入力される情報を含む。図6を参照して、第3プロンプトの構成について説明する。図6に示すように、ステップS106において生成される第3プロンプトは、「インストラクション」として第3インストラクションデータと、「負例のサンプル」として第3関係知識データと、「負例ではないサンプル」として第1関係知識データと、「クエリ」として第2関係知識データと、を含む。

[0064] 例えば、判定部16Aは、第3関係知識データセットから一部または全部の第3関係知識データを抽出し、抽出した各第3関係知識データを負例のサンプルとして第3プロンプトに含める。また、判定部16Aは、第1関係知識データセットから一部または全部の第1関係知識データを抽出し、抽出した各第1関係知識データを、負例ではないサンプルとして第3プロンプトに含める。また、例えば、判定部16Aは、第2関係知識データセットに含まれる各第2関係知識データを「クエリ」として第3プロンプトに含める。また、例えば、判定部16Aは、記憶部120に記憶された第3インストラクションデータを「インストラクション」として第3プロンプトに含める。第3インストラクションデータは、第3プロンプトにより生成したい情報を言語モデル3Aに教示する内容を示す自然言語文であり、テンプレートとして事前に記憶部120に記憶されている。

[0065] 図14は、第3プロンプトの具体例を説明する図である。図14の例では、第3インストラクションデータは、当該第3プロンプトにより生成したい情報が第1項目「INDUSTRY」、第2項目「SENTENCE1」、第3項目「SENTENCE2」、および第4項目「Label」の各項目を含むこと、並びに、各項目の定義を説明する自然言語文を含む。この例では、第4項目「Label」の定義として、「e」が異なる関係であることを示し、「x」が異なる関係でないことを示すことを説明する自然言語文が、第3インストラクションデータに含まれている。

[0066] また、図14の例では、第3プロンプトに含まれるサンプルは、負例のサンプルSMP31、負例ではないサンプルSMP32を含む。負例のサンプルSMP31は、第1項目～第3項目として図13に示した第3関係知識データD31を含み、第4項目「Label」として「e」を含む。ここでは「e」は「異なる関係」を示すラベルである。負例ではないサンプルSMP32は、第1項目～第3項目として図7に示した第1関係知識データD12を含み、第4項目「Label」として「x」を含む。ここでは「x」は「異なる関係ではない」ことを示すラベルである。

[0067] ここで、負例のサンプルSMP31を生成するための第3関係知識データD31は、判定部16Aによって第3関係知識データセットから抽出されたものであってよい。また、負例ではないサンプルSMP32を生成するための第1関係知識データD12は、判定部16Aによって第1関係知識データセットから抽出されたものであってよい。当該抽出処理は、ランダムに行われてもよいし、所定の条件を満たすものが抽出されてもよい。例えば、所定の条件とは、スコア、生成日時、過去の抽出履歴、等に基づく条件であってもよい。

[0068] なお、ステップS106における第3プロンプトの生成処理は、第2関係知識データセットに含まれる各第2関係知識データについて個別に実行されてもよいし、一部または複数の第2関係知識データについて共通して1回実行されてもよい。個別に実行される場合、各第2関係知識データについて生成される第3プロンプトに含まれるサンプルは、互いに異なるものとなり得る。これにより、サンプルの内容的な偏りを軽減でき、より精度よく負例であるか否かを判定可能となる。なお、図14の例では、第3プロンプトに含まれる1つの負例のサンプルSMP31を示しているが、第3プロンプトに含まれる負例のサンプルの個数は複数であってもよい。また、第3プロンプトに含まれる1つの負例ではないサンプルSMP32を示しているが、第3プロンプトに含まれる負例ではないサンプルの個数は複数であってもよい。

[0069] また、図14の例では、第3プロンプトに含まれるクエリは、第1項目～

第3項目として第2関係知識データを含むが、第4項目「L a b e l」として項目名のみを含み当該項目名に続く自然言語文を含まない。このようなクエリは、クエリに含まれる第2関係知識データに対応する「L a b e l」を示す自然言語文の生成要求を示している。

[0070] 次に、ステップS 1 0 7において、判定部1 6 Aは、生成した第3プロンプトを言語モデル3 Aに入力することにより、第2関係知識データセットに含まれる各第2関係知識データが負例であるか否かを判定する。例えば、判定部1 6 Aは、「L a b e l」として「e」が生成された第2関係知識データを負例（すなわち、「異なる関係」を有する）と判定する。また、例えば、判定部1 6 Aは、「L a b e l」として「x」が生成された第2関係知識データを負例ではない（すなわち、異なる関係を有していない）と判定する。

[0071] 次に、ステップS 1 0 8において、判定部1 6 Aは、ステップS 1 0 7で負例ではないと判定された第2関係知識データについて、さらにユーザからの入力に基づいて、正例であるか負例であるかを判定する。なお、ユーザは、例えば、負例ではないと判定された第2関係知識データについて、特定の関係を有する正例、異なる関係を有する負例、および、それ以外、の3つの何れであるかを示すアノテーション情報を入力してもよい。これにより、言語モデル3 Aにより「負例でない」と判定された第2関係知識データを全て正例とする場合と比べて、より精度よく、「特定の関係を有する正例」を特定することができる。

[0072] このようにして、図6のステップS 1 0 7～S 1 0 8に示すように、第2関係知識データセットに含まれる各第2関係知識データは、正例または負例に分類される。正例に分類された第2関係知識データは、言語モデル3 Aにより負例ではないと判定され、かつ、ユーザによって正例であると選択されたものである。負例に分類された第2関係知識データは、言語モデル3 Aにより負例であると判定されたものを含む。また、負例に分類された第2関係知識データは、言語モデル3 Aにより負例ではないと判定されたが、ユーザ

によって負例であると選択されたものを含み得る。

[0073] (ステップS109～S111:更新)

再び図4、図6を参照して、ステップS109～S111について説明する。ステップS109～S111は、第1関係知識データセット、第3関係知識データセット、および関係知識データベース2Aを更新するステップである。図6のステップS109～S111に示すように、正例に分類された第2関係知識データは、第1関係知識データセットおよび関係知識データベース2Aを更新するために用いられる。また、負例に分類された第2関係知識データは、第3関係知識データセットを更新するために用いられる。

[0074] ステップS109において、更新部13Aは、第2関係知識データセットのうち、負例に分類された第2関係知識データを用いて、第3関係知識データセットを更新する。「第3関係知識データセットを更新」するとは、例えば、負例に分類された第2関係知識データを新たな第3関係知識データとして、第3関係知識データセットに追加することであってもよい。また、「第3関係知識データセットを更新」するとは、例えば、負例に分類された第2関係知識データを新たな第3関係知識データとして、第3関係知識データセットの一部または全部を置き換えることであってもよい。

[0075] また、言語モデル3Aが、第3プロンプトに応じて、各第2関係知識データについて生成した第4項目「Label」の適切さを示すスコアを出力している場合がある。この場合、更新部13Aは、負例に分類された第2関係知識データのうち、スコアが所定の条件を満たす（例えば、スコアが閾値以上である、スコアが上位から所定順位以内、所定割合以内である、等）ものを用いて、第3関係知識データセットを更新してもよい。また、更新部13Aは、負例に分類された第2関係知識データのうち、ユーザによって選択されたものを用いて第3関係知識データセットを更新してもよい。

[0076] ステップS109を実行することにより、次に情報処理方法S1Aを繰り返す際に参照される第1関係知識データセットが更新されるので、第1関係知識データセットに基づき生成される第2プロンプトの精度が向上する。そ

の結果、生成される第2関係知識データの正例としての精度が向上する。

[0077] ステップS110において、更新部13Aは、第2関係知識データセットのうち、正例に分類された第2関係知識データを用いて、第1関係知識データセットを更新する。「第1関係知識データセットを更新」するとは、例えば、正例に分類された第2関係知識データを新たな第1関係知識データとして、第1関係知識データセットに追加することであってもよい。また、「第1関係知識データセットを更新」するとは、例えば、正例に分類された第2関係知識データを新たな第1関係知識データとして、第1関係知識データセットの一部または全部を置き換えることであってもよい。なお、更新された第1関係知識データセットは、記憶部120に記憶されてもよい。

[0078] また、生成部12Aが、第2プロンプトに応じて、各コンテキストに応じて生成した第2項目「STEER」（すなわち、第2関係知識データ）の適切さを示すスコアを出力している場合がある。この場合、更新部13Aは、正例に分類された第2関係知識データのうち、当該スコアが所定の条件を満たす（例えば、スコアが閾値以上である、スコアが上位から所定順位以内、所定割合以内である、等）ものを用いて、第1関係知識データセットを更新してもよい。また、更新部13Aは、正例に分類された第2関係知識データのうち、ユーザによって選択されたものを用いて第1関係知識データセットを更新してもよい。

[0079] ステップS109を実行することにより、次に情報処理方法S1Aを繰り返す際に参照される第3関係知識データセットが更新されるので、第3関係知識データセットに基づき生成される第3プロンプトの精度が向上する。その結果、第2関係知識データについての負例の判定精度が向上する。これにより、第2関係知識データセットから負例を精度よく除外することができる。

[0080] ステップS111において、更新部13Aは、第2関係知識データセットのうち、正例に分類された第2関係知識データを用いて、関係知識データベース2Aを更新する。例えば、更新部13Aは、正例に分類された第2関係

知識データを新たな「特定の関係を有する関係知識データ」として関係知識データベース2 Aに追加する。

[0081] そして、情報処理装置1 Aは、情報処理方法S 1 Aを繰り返す。次に繰り返される情報処理方法S 1 Aでは、ステップS 1 0 1において、取得部1 1 Aは、更新された第1関係知識データセットを取得する。また、ステップS 1 0 2、S 1 0 4において、生成部1 2 Aは、更新された第1関係知識データセットを用いて第1プロンプト、第2プロンプトを生成する。これにより、ステップS 1 0 5において、更新された第1関係知識データセットを事例として参照した新たな第2関係知識データセットが生成される。また、ステップS 1 0 6において、判定部1 6 Aは、更新された第1関係知識データセットおよび更新された第3関係知識データセットを用いて第3プロンプトを生成する。

[0082] このようにして情報処理方法S 1 Aを繰り返すことにより、指定された各コンテキストに応じて正例としての精度がより高い第2関係知識データセットを生成できるとともに、そのうち負例をより精度よく判定して除外することができる。その結果、関係知識データベース2 Aが精度よく更新される。

[0083] <本例示的实施形態の効果>

以上説明したように、本例示的实施形態は、更新された第1関係知識データセットを事例として参照することにより、新たな第2関係知識データセットを生成する、との構成が採用されている。このため、本例示的实施形態によれば、特定の関係を有するものとしてより精度の高い第1関係知識データセットを事例として参照できるので、関係知識データベース2 Aに追加すべき特定の関係を有する関係知識データを、指定されるコンテキストに応じて精度よく生成することができる。

[0084] また、本例示的实施形態により関係知識データベース2 Aに蓄積される関係知識データは、指定されるコンテキストにおける判断、分析等に活用することができる。その場合、そのような判断、分析等を精度よく行うことができる。例えば、特定の関係として「因果関係」を適用し、コンテキストとし

て「産業の種別」を適用する場合、関係知識データベース 2 A に蓄積される関係知識データは、指定された業界における経営判断、業界分析等に活用することができる。その場合、当該業界における変化／変革を精度よく把握するとともに、それに伴う将来変化を精度よく予測することができる。

[0085] また、本例示的实施形態は、第 2 関係知識データセットのうち、言語モデル 3 A が算出したスコアに基づく条件を満たす第 2 関係知識データを、特定の関係を有する第 2 関係知識データとして用いて、第 1 関係知識データセットを更新する、という構成が採用されている。このため、本例示的实施形態によれば、スコアに基づき特定の関係を有する可能性が高い第 2 関係知識データを用いて第 1 関係知識データセットを更新できるので、特定の関係を有する事例として参照される第 1 関係知識データセットの精度を向上させることができる。

[0086] また、本例示的实施形態は、第 2 関係知識データセットのうち、ユーザによって選択された第 2 関係知識データを、特定の関係を有する第 2 関係知識データとして用いて、第 1 関係知識データセットを更新する、という構成が採用されている。このため、本例示的实施形態によれば、特定の関係を有するものとしてユーザが選択した第 2 関係知識データを用いて第 1 関係知識データセットを更新できるので、特定の関係を有する事例として参照される第 1 関係知識データセットの精度を向上させることができる。

[0087] また、本例示的实施形態は、言語モデル 3 A を用いて、特定の関係とは異なる関係を有する複数の知識を示す第 3 関係知識データによって構成される第 3 関係知識データセットを事例として参照することにより、第 2 関係知識データが異なる関係を有するか否かを判定し、第 2 関係知識データセットのうち、異なる関係を有していないと判定された第 2 関係知識データを、特定の関係を有する第 2 関係知識データとして用いて、第 1 関係知識データセットを更新する、との構成が採用されている。このため、本例示的实施形態によれば、言語モデル 3 A を用いて生成した第 2 関係知識データセットから負例を除外するために同一の言語モデル 3 A を用いるので、正確性を担保でき

る。また、言語モデル 3 A を用いて生成した第 2 関係知識データセットから負例を除外するために異なるモデルを用いる必要がないので、より容易に負例を除外することができる。

[0088] また、本例示的实施形態は、第 2 関係知識データセットのうち、異なる関係を有していると判定された第 2 関係知識データを用いて、第 3 関係知識データセットを更新する、との構成が採用されている。このため、本例示的实施形態によれば、言語モデル 3 A を用いて生成した第 2 関係知識データセットから負例を除外するために事例として参照される、特定の関係とは異なる関係を有する第 3 関係知識データセットの精度が向上するので、結果として、関係知識データベース 2 A に追加すべき特定の関係を有する関係知識データを精度よく抽出することができる。

[0089] また、本例示的实施形態では、第 1 関係知識データは、前件の知識を示す第 1 前件データ、前件の知識に対して特定の関係を有する後件の知識を示す第 1 後件データを含み、言語モデル 3 A を用いて、第 1 前件データを事例として参照することにより、コンテキストに応じた第 2 前件データを生成するとともに、第 1 関係知識データを事例として参照することにより、生成した第 2 前件データに応じた第 2 後件データを生成することにより、第 2 前件データおよび第 2 後件データを含む第 2 関係知識データを生成する、との構成が採用されている。このため、本例示的实施形態によれば、指定されたコンテキストに応じた第 2 前件データをまず作成してから、当該第 2 前件データに応じた第 2 後件データを生成するので、より精度よく、指定されたコンテキストに応じた第 2 関係知識データを生成することができる。

[0090] [変形例 1]

上述した例示的实施形態 2 は、特定の関係とは異なる関係として、複数の関係を適用するよう変形可能である。本変形例では、異なる関係として第 1 の関係および第 2 の関係を適用する例について説明する。

[0091] 本変形例では、第 3 関係知識データセットは、特定の関係とは異なる第 1 の関係を有する第 3 関係知識データと、特定の関係および第 1 の関係の何れ

とも異なる第2の関係性を有する第3関係知識データと、を含む。例えば、「特定の関係」として「因果関係」を適用する場合、「第1の関係」として「同義関係」を適用し、第2の関係として「含意関係」を適用してもよい。ただし、第1の関係および第2の関係は、これらの例に限られない。

[0092] 本変形例では、判定部16Aは、第2関係知識データセットに含まれる各第2関係知識データが、第1の関係および第2の関係の何れを有しているか否かを判定する。以降では、第1の関係を有する複数の知識を第1負例とも記載し、第2の関係を有する複数の知識を第2負例とも記載する。

[0093] 例えば、本変形例では、第3プロンプトに含まれるサンプルが次のように変形される。具体的には、判定部16Aは、第3関係知識データセットから、第1の関係を有する一部または全部の第3関係知識データを抽出し、抽出した各第3関係知識データを第1負例のサンプルとして第3プロンプトに含める。また、判定部16Aは、第3関係知識データセットから、第2の関係を有する一部または全部の第3関係知識データを抽出し、抽出した各第3関係知識データを第2負例のサンプルとして第3プロンプトに含める。

[0094] また、例えば、第3プロンプトに含める「インストラクション」としての第3インストラクションデータにおいて、第4項目「Label」の定義は、次のように変形される。例えば、第4項目「Label」の定義は、「e」が第1負例であることを示し、「a」が第2負例であることを示し、「x」が第1負例および第2負例の何れでもないことを示すことを説明する自然言語文に変形される。

[0095] このようにして、判定部16Aは、第3プロンプトを言語モデル3Aに入力することにより、第2関係知識データセットに含まれる各第2関係知識データが第1負例、第2負例、およびその他の何れであるか否かを判定する。また、判定部16Aは、言語モデル3Aにより第1負例および第2負例の何れでもないと判定した第2関係知識データについて、さらに、ユーザの入力に基づいて「正例」、「第1負例」、「第2負例」、および「その他」の何れであるかを判定してもよい。

[0096] また、本変形例において、更新部 13A は、第 2 関係知識データセットのうち、第 1 負例および第 2 負例の何れでもないと判定された第 2 関係知識データを、特定の関係を有する第 2 関係知識データとして用いて、第 1 関係知識データセットを更新する。具体的には、更新部 13A は、第 1 負例および第 2 負例の何れでもないと判定された第 2 関係知識データのうち、ユーザによって「正例」として選択された第 2 関係知識データを用いて、第 1 関係知識データセットを更新する。また、更新部 13A は、第 1 負例および第 2 負例の何れでもないと判定された第 2 関係知識データのうち、ユーザによって「第 1 負例」として選択された第 2 関係知識データを用いて、第 3 関係知識データセットを更新する。また、更新部 13A は、第 1 負例および第 2 負例の何れでもないと判定された第 2 関係知識データのうち、ユーザによって「第 2 負例」として選択された第 2 関係知識データを用いて、第 3 関係知識データセットを更新する。

[0097] 本変形例によれば、第 2 関係知識データセットから、特定の関係を有していないとして除外する負例として複数種類を判定するので、特定の関係を有していない負例をより精度よく除外することができる。なお、ここでは、負例が 2 種類である例について説明したが、負例は 3 種類以上であってもよい。

[0098] [変形例 2]

上述した変形例 2 は、さらに、判定部 16A が、第 1 負例であるか否かおよび第 2 負例であるか否かを段階的に判定するように変形可能である。

[0099] 本変形例においては、判定部 16A は、第 1 の第 3 プロンプトと、第 2 の第 3 プロンプトとを用いる。第 1 の第 3 プロンプトは、例示的实施形態 2 における第 3 プロンプトの説明において「異なる関係」を「第 1 の関係」に読み替えることにより同様に説明される。また、第 2 の第 3 プロンプトは、例示的实施形態における第 3 プロンプトの説明において「異なる関係」を「第 2 の関係」に読み替えることにより同様に説明される。

[0100] 判定部 16A は、まず、各第 2 関係知識データについて、第 1 の第 3 プロ

ンプトを言語モデル 3 A に入力することにより、第 1 負例であるか否かを判定する。また、判定部 1 6 A は、第 1 負例ではないと判定された第 2 関係知識データについて、第 2 の第 3 プロンプトを言語モデル 3 A に入力することにより、第 2 負例であるか否かを判定する。これにより、判定部 1 6 A は、各第 2 関係知識データを、第 1 負例、第 2 負例、および、その他の何れであるかを判定することができる。

[0101] 本変形例によれば、第 2 関係知識データセットから、特定の関係を有していないとして除外する複数種類の負例を段階的に判定するので、1 つの段階では 2 値判定を行えばよく、より精度よく負例の判定を行うことができる。

[0102] なお、ここでは、2 段階に分けて判定を行う例について説明したが、負例の種類が 3 種類以上である場合に、判定の段階数も 3 以上であってもよい。また、負例の種類数と段階数とは必ずしも一致していなくてもよい。例えば、判定部 1 6 A は、各第 2 関係知識データについて、ある判定段階では第 1 負例であるか否かを判定し、他の判定段階では第 2 負例および第 3 負例の何れであるか否かを判定してもよい。

[0103] [その他の変形例]

上述した例示的实施形態 2 では、第 1 関係知識データセットを更新するために用いる正例の第 2 関係知識データを特定するために、「ユーザによる選択」、「スコアに基づく条件」、および「言語モデル 3 A による負例の除外」の全てを組み合わせる例について説明した。ただし、これらの一部を適用し、他の一部を適用せずに、正例の第 2 関係知識データを特定するよう変形することも可能である。

[0104] 上述した例示的实施形態 2 では、第 3 関係知識データセットを更新するために用いる負例の第 2 関係知識データを特定するために、「ユーザによる選択」、「スコアに基づく条件」、および「言語モデル 3 A による負例の判定」の全てを組み合わせる例について説明した。ただし、これらの一部を適用し、他の一部を適用せずに、負例の第 2 関係知識データを特定するよう変形することも可能である。

[0105] 上述した例示的实施形態2では、第1関係知識データ、第2関係知識データ、および第3関係知識データは、それぞれ、2つの知識の関係を示すデータであるとして説明した。ただし、これらの関係知識データは、3つ以上の知識の関係を示すデータであってもよい。

[0106] [ソフトウェアによる実現例]

情報処理装置1、1A（以下、それぞれを当該装置と記載）の一部又は全部の機能は、集積回路（ICチップ）等のハードウェアによって実現してもよいし、ソフトウェアによって実現してもよい。

[0107] 後者の場合、当該装置は、例えば、各機能を実現するソフトウェアであるプログラムの命令を実行するコンピュータによって実現される。このようなコンピュータの一例（以下、コンピュータCと記載する）を図15に示す。コンピュータCは、少なくとも1つのプロセッサC1と、少なくとも1つのメモリC2と、を備えている。メモリC2には、コンピュータCを当該装置として動作させるためのプログラムPが記録されている。コンピュータCにおいて、プロセッサC1は、プログラムPをメモリC2から読み取って実行することにより、当該装置の各機能が実現される。

[0108] プロセッサC1としては、例えば、CPU（Central Processing Unit）、GPU（Graphic Processing Unit）、DSP（Digital Signal Processor）、MPU（Micro Processing Unit）、FPU（Floating point number Processing Unit）、PPU（Physics Processing Unit）、TPU（Tensor Processing Unit）、量子プロセッサ、マイクロコントローラ、又は、これらの組み合わせなどを用いることができる。メモリC2としては、例えば、フラッシュメモリ、HDD（Hard Disk Drive）、SSD（Solid State Drive）、又は、これらの組み合わせなどを用いることができる。

[0109] なお、コンピュータCは、プログラムPを実行時に展開したり、各種データを一時的に記憶したりするためのRAM（Random Access Memory）を更に備えていてもよい。また、コンピュータCは、他の装置との間でデータを送受信するための通信インタフェースを更に備えていてもよい。また、コンピ

ユータCは、キーボードやマウス、ディスプレイやプリンタなどの入出力機器を接続するための入出力インタフェースを更に備えていてもよい。

[0110] また、プログラムPは、コンピュータCが読み取り可能な、一時的でない有形の記録媒体Mに記録することができる。このような記録媒体Mとしては、例えば、テープ、ディスク、カード、半導体メモリ、又はプログラマブルな論理回路などを用いることができる。コンピュータCは、このような記録媒体Mを介してプログラムPを取得することができる。また、プログラムPは、伝送媒体を介して伝送することができる。このような伝送媒体としては、例えば、通信ネットワーク、又は放送波などを用いることができる。コンピュータCは、このような伝送媒体を介してプログラムPを取得することもできる。

[0111] [付記事項1]

本発明は、上述した実施形態に限定されるものでなく、請求項に示した範囲で種々の変更が可能である。例えば、上述した実施形態に開示された技術的手段を適宜組み合わせ得られる実施形態についても、本発明の技術的範囲に含まれる。

[0112] [付記事項2]

上述した実施形態の一部又は全部は、以下のようにも記載され得る。ただし、本発明は、以下の記載する態様に限定されるものではない。

[0113] (付記1)

特定の関係を有する複数の知識を示す第1関係知識データによって構成される第1関係知識データセット、および、知識のコンテキストを指定する情報を取得する取得手段と、

言語モデルを用いて、前記第1関係知識データセットを事例として参照することにより、前記コンテキストに応じた第2関係知識データによって構成される第2関係知識データセットを生成する生成手段と、

前記第2関係知識データセットのうち、前記特定の関係を有する第2関係知識データを用いて、前記第1関係知識データセットを更新する更新手段と

、
を含む、情報処理装置。

[0114] (付記2)

前記生成手段は、前記更新された第1関係知識データセットを事例として参照することにより、新たな前記第2関係知識データセットを生成する、
付記1に記載の情報処理装置。

[0115] (付記3)

前記更新手段は、前記第2関係知識データセットのうち、前記言語モデルが算出したスコアに基づく条件を満たす第2関係知識データを、前記特定の関係性を有する第2関係知識データとして用いて、前記第1関係知識データセットを更新する、
付記1または2に記載の情報処理装置。

[0116] (付記4)

前記更新手段は、前記第2関係知識データセットのうち、ユーザによって選択された第2関係知識データを、前記特定の関係性を有する第2関係知識データとして用いて、前記第1関係知識データセットを更新する、
付記1から3の何れか1つに記載の情報処理装置。

[0117] (付記5)

前記言語モデルを用いて、前記特定の関係とは異なる関係を有する複数の知識を示す第3関係知識データによって構成される第3関係知識データセットを事例として参照することにより、前記第2関係知識データが前記異なる関係を有するか否かを判定する判定手段、
をさらに含み、

前記更新手段は、前記第2関係知識データセットのうち、前記異なる関係を有していないと判定された第2関係知識データを、前記特定の関係性を有する第2関係知識データとして用いて、前記第1関係知識データセットを更新する、

付記1から4の何れか1つに記載の情報処理装置。

[0118] (付記6)

前記更新手段は、前記第2関係知識データセットのうち、前記異なる関係を有していると判定された前記第2関係知識データを用いて、前記第3関係知識データセットを更新する、

付記5に記載の情報処理装置。

[0119] (付記7)

前記第3関係知識データセットは、前記特定の関係とは異なる第1の関係を有する第3関係知識データと、前記特定の関係および前記第1の関係の何れとも異なる第2の関係を有する第3関係知識データと、を含み、

前記判定手段は、前記第2関係知識データが前記第1の関係および前記第2の関係の何れを有しているか否かを判定し、

前記更新手段は、前記第2関係知識データセットのうち、前記第1の関係および前記第2の関係を何れも有していないと判定した前記第2関係知識データを、前記特定の関係を有する第2関係知識データとして用いて、前記第1関係知識データセットを更新する、

付記5または6に記載の情報処理装置。

[0120] (付記8)

前記第1関係知識データは、前件の知識を示す第1前件データ、前記前件の知識に対して前記特定の関係を有する後件の知識を示す第1後件データを含み、

前記生成手段は、前記言語モデルを用いて、

前記第1前件データを事例として参照することにより、前記コンテキストに応じた第2前件データを生成する第1生成手段と、

前記第1関係知識データを事例として参照することにより、前記第2前件データに応じた第2後件データを生成する第2生成手段と、を含み、

前記第2前件データおよび前記第2後件データを含む前記第2関係知識データを生成する、

付記1から7の何れか1つに記載の情報処理装置。

[0121] (付記9)

1以上のプロセッサが実行する情報処理方法であって、
特定の関係を有する複数の知識を示す第1関係知識データによって構成される第1関係知識データセット、および、知識のコンテキストを指定する情報を取得することと、
言語モデルを用いて、前記第1関係知識データセットを事例として参照することにより、前記コンテキストに応じた第2関係知識データによって構成される第2関係知識データセットを生成することと、
前記第2関係知識データセットのうち、前記特定の関係を有する第2関係知識データを用いて、前記第1関係知識データセットを更新することと、
を含む、情報処理方法。

[0122] (付記10)

コンピュータを、
特定の関係を有する複数の知識を示す第1関係知識データによって構成される第1関係知識データセット、および、知識のコンテキストを指定する情報を取得する取得手段と、
言語モデルを用いて、前記第1関係知識データセットを事例として参照することにより、前記コンテキストに応じた第2関係知識データによって構成される第2関係知識データセットを生成する生成手段と、
前記第2関係知識データセットのうち、前記特定の関係を有する第2関係知識データを用いて、前記第1関係知識データセットを更新する更新手段と、
、
として機能させるプログラム。

[0123] [付記事項3]

上述した実施形態の一部又は全部は、更に、以下のように表現することもできる。

[0124] 少なくとも1つのプロセッサを備え、前記プロセッサは、

特定の関係を有する複数の知識を示す第1関係知識データによって構成さ

れる第1関係知識データセット、および、知識のコンテキストを指定する情報を取得する取得処理と、

言語モデルを用いて、前記第1関係知識データセットを事例として参照することにより、前記コンテキストに応じた第2関係知識データによって構成される第2関係知識データセットを生成する生成処理と、

前記第2関係知識データセットのうち、前記特定の関係性を有する第2関係知識データを用いて、前記第1関係知識データセットを更新する更新処理と、

を実行する情報処理装置。

[0125] なお、この情報処理装置は、更にメモリを備えていてもよく、このメモリには、前記取得処理と、前記生成処理と、前記更新処理とを前記プロセッサに実行させるためのプログラムが記憶されていてもよい。また、このプログラムは、コンピュータ読み取り可能な一時的でない有形の記録媒体に記録されていてもよい。

符号の説明

- [0126] 1、1 A 情報処理装置
2 A 関係知識データベース
3 A 言語モデル
1 1、1 1 A 取得部
1 2、1 2 A 生成部
1 3、1 3 A 更新部
1 4 A 第1生成部
1 5 A 第2生成部
1 6 A 判定部
1 0 0 情報処理システム
1 1 0 制御部
1 2 0 記憶部
C 1 プロセッサ

C 2 メモリ

請求の範囲

- [請求項1] 特定の関係を有する複数の知識を示す第1関係知識データによって構成される第1関係知識データセット、および、知識のコンテキストを指定する情報を取得する取得手段と、
- 言語モデルを用いて、前記第1関係知識データセットを事例として参照することにより、前記コンテキストに応じた第2関係知識データによって構成される第2関係知識データセットを生成する生成手段と、
- 前記第2関係知識データセットのうち、前記特定の関係を有する第2関係知識データを用いて、前記第1関係知識データセットを更新する更新手段と、
- を含む、情報処理装置。
- [請求項2] 前記生成手段は、前記更新された第1関係知識データセットを事例として参照することにより、新たな前記第2関係知識データセットを生成する、
- 請求項1に記載の情報処理装置。
- [請求項3] 前記更新手段は、前記第2関係知識データセットのうち、前記言語モデルが算出したスコアに基づく条件を満たす第2関係知識データを、前記特定の関係を有する第2関係知識データとして用いて、前記第1関係知識データセットを更新する、
- 請求項1または2に記載の情報処理装置。
- [請求項4] 前記更新手段は、前記第2関係知識データセットのうち、ユーザによって選択された第2関係知識データを、前記特定の関係を有する第2関係知識データとして用いて、前記第1関係知識データセットを更新する、
- 請求項1から3の何れか1項に記載の情報処理装置。
- [請求項5] 前記言語モデルを用いて、前記特定の関係とは異なる関係を有する複数の知識を示す第3関係知識データによって構成される第3関係知

識データセットを事例として参照することにより、前記第2関係知識データが前記異なる関係を有するか否かを判定する判定手段、
をさらに含み、

前記更新手段は、前記第2関係知識データセットのうち、前記異なる関係を有していないと判定された第2関係知識データを、前記特定の関係を有する第2関係知識データとして用いて、前記第1関係知識データセットを更新する、

請求項1から4の何れか1項に記載の情報処理装置。

[請求項6]

前記更新手段は、前記第2関係知識データセットのうち、前記異なる関係を有していると判定された前記第2関係知識データを用いて、前記第3関係知識データセットを更新する、

請求項5に記載の情報処理装置。

[請求項7]

前記第3関係知識データセットは、前記特定の関係とは異なる第1の関係を有する第3関係知識データと、前記特定の関係および前記第1の関係を何れとも異なる第2の関係を有する第3関係知識データと、
を含み、

前記判定手段は、前記第2関係知識データが前記第1の関係および前記第2の関係を何れを有しているか否かを判定し、

前記更新手段は、前記第2関係知識データセットのうち、前記第1の関係および前記第2の関係を何れも有していないと判定した前記第2関係知識データを、前記特定の関係を有する第2関係知識データとして用いて、前記第1関係知識データセットを更新する、

請求項5または6に記載の情報処理装置。

[請求項8]

前記第1関係知識データは、前件の知識を示す第1前件データ、前記前件の知識に対して前記特定の関係を有する後件の知識を示す第1後件データを含み、

前記生成手段は、前記言語モデルを用いて、

前記第1前件データを事例として参照することにより、前記コン

テキストに応じた第2前件データを生成する第1生成手段と、

前記第1関係知識データを事例として参照することにより、前記第2前件データに応じた第2後件データを生成する第2生成手段と、
を含み、

前記第2前件データおよび前記第2後件データを含む前記第2関係知識データを生成する、

請求項1から7の何れか1項に記載の情報処理装置。

[請求項9]

1以上のプロセッサが実行する情報処理方法であって、

特定の関係を有する複数の知識を示す第1関係知識データによって構成される第1関係知識データセット、および、知識のコンテキストを指定する情報を取得することと、

言語モデルを用いて、前記第1関係知識データセットを事例として参照することにより、前記コンテキストに応じた第2関係知識データによって構成される第2関係知識データセットを生成することと、

前記第2関係知識データセットのうち、前記特定の関係を有する第2関係知識データを用いて、前記第1関係知識データセットを更新することと、

を含む、情報処理方法。

[請求項10]

コンピュータを、

特定の関係を有する複数の知識を示す第1関係知識データによって構成される第1関係知識データセット、および、知識のコンテキストを指定する情報を取得する取得手段と、

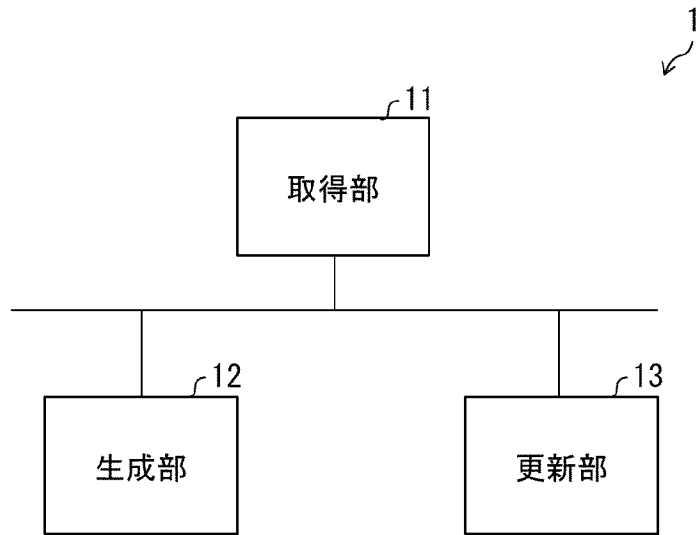
言語モデルを用いて、前記第1関係知識データセットを事例として参照することにより、前記コンテキストに応じた第2関係知識データによって構成される第2関係知識データセットを生成する生成手段と、
、

前記第2関係知識データセットのうち、前記特定の関係を有する第2関係知識データを用いて、前記第1関係知識データセットを更新す

る更新手段と、
として機能させるプログラム。

[図1]

図 1



[図2]

図 2

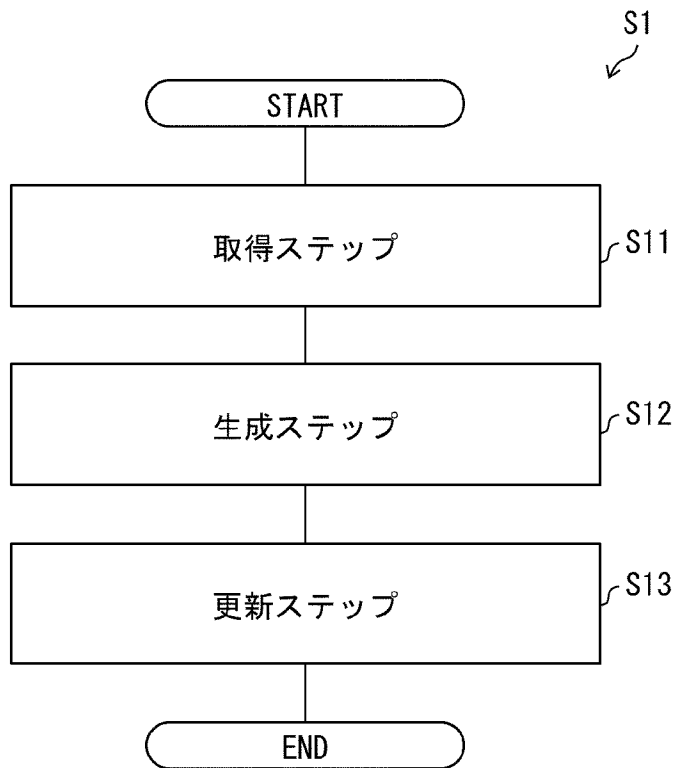


図3

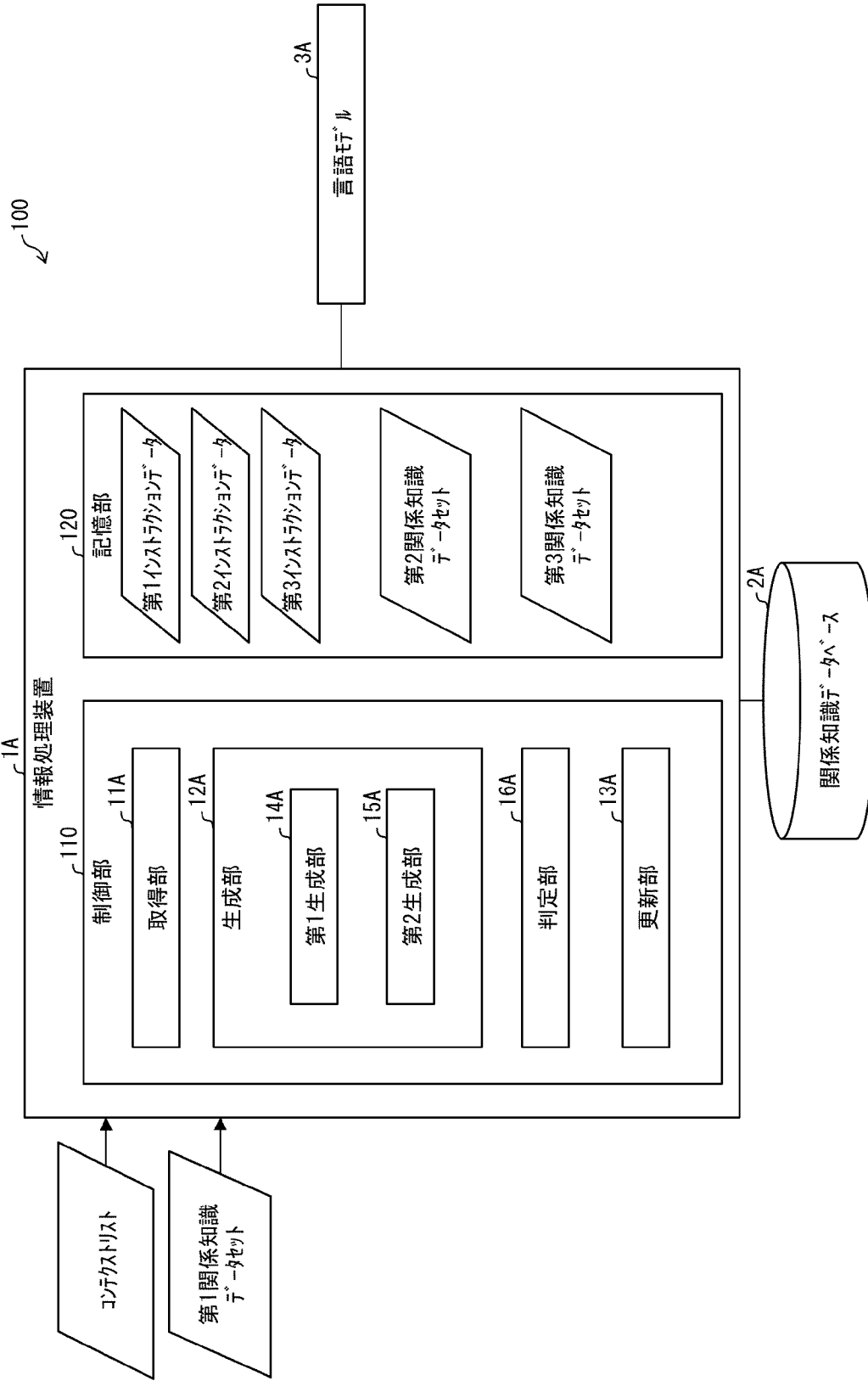
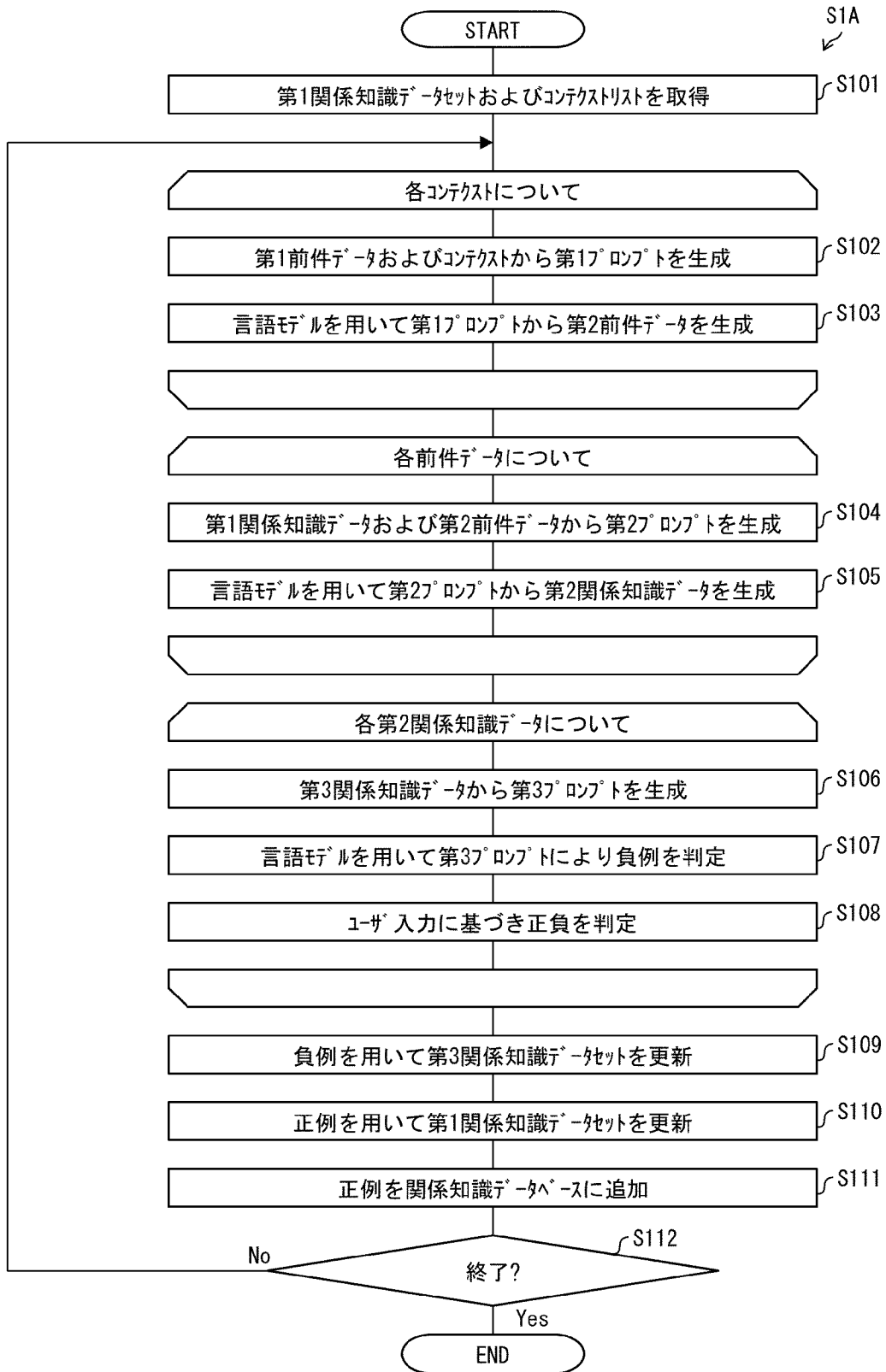


図3

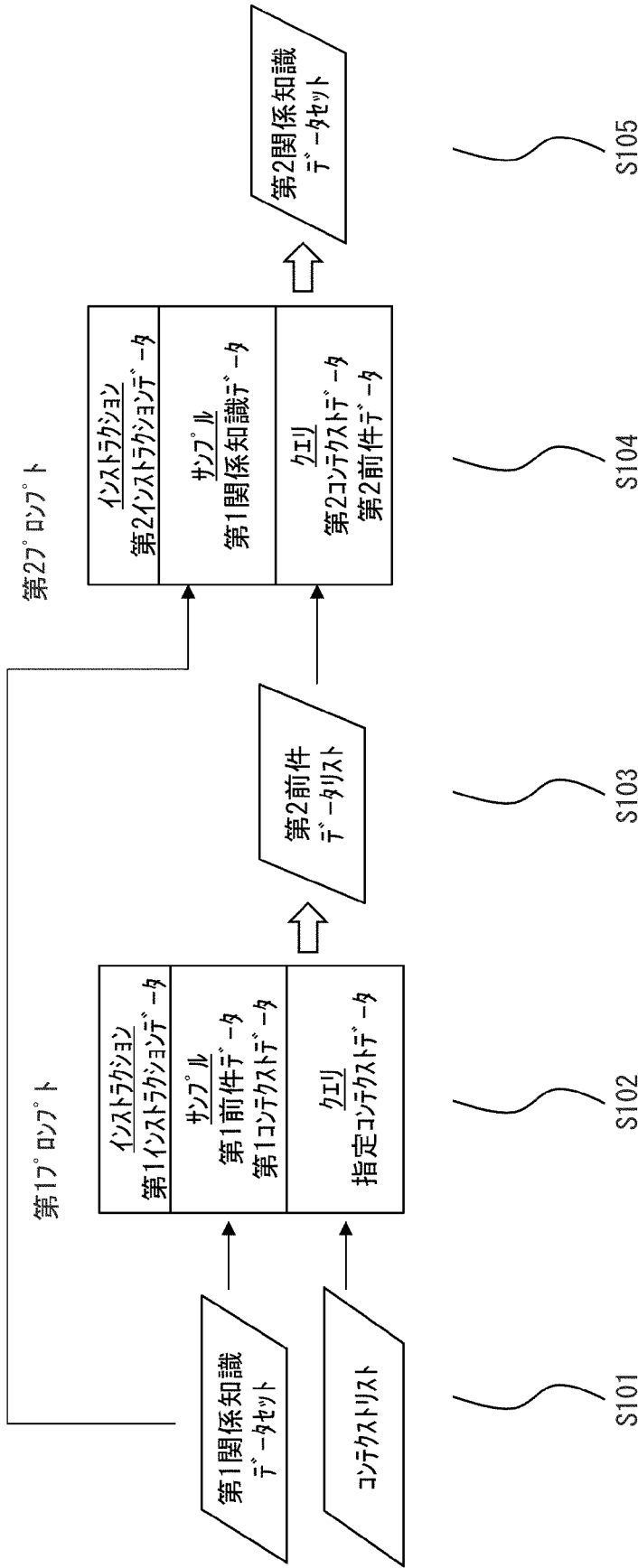
[図4]

図 4



[図5]

図 5



[図6]

図 6

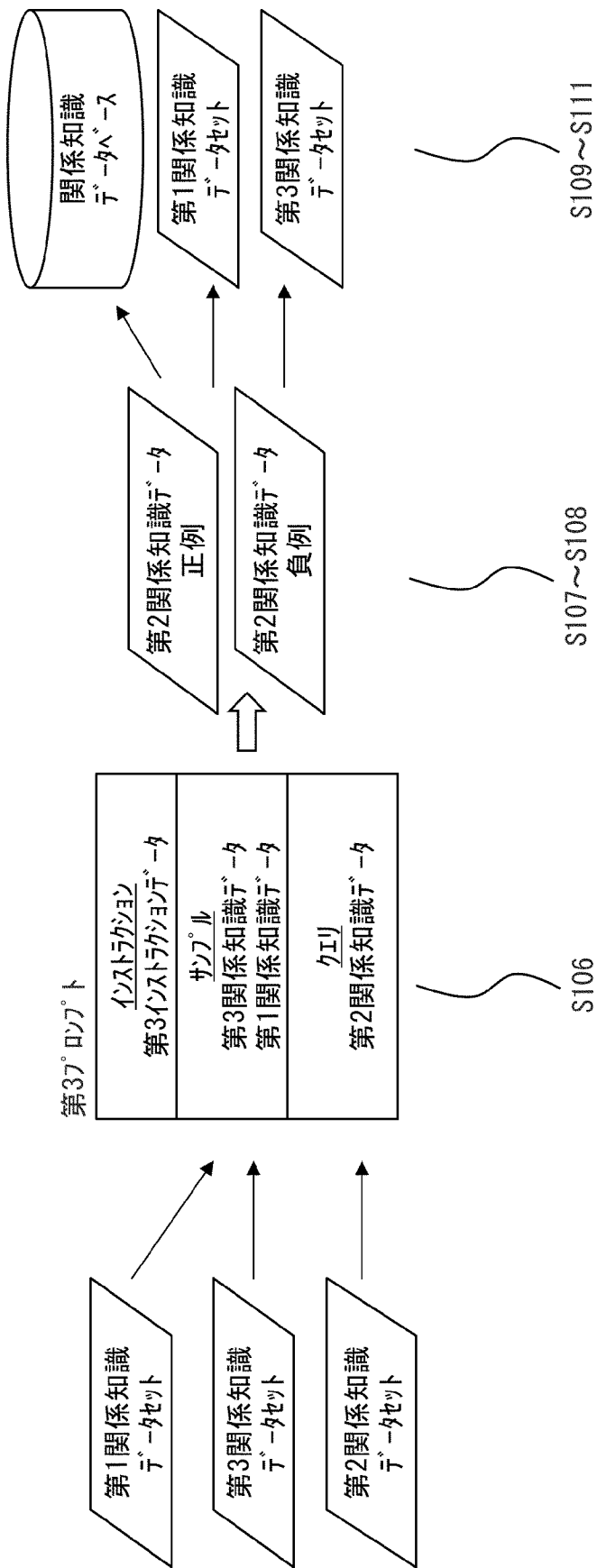
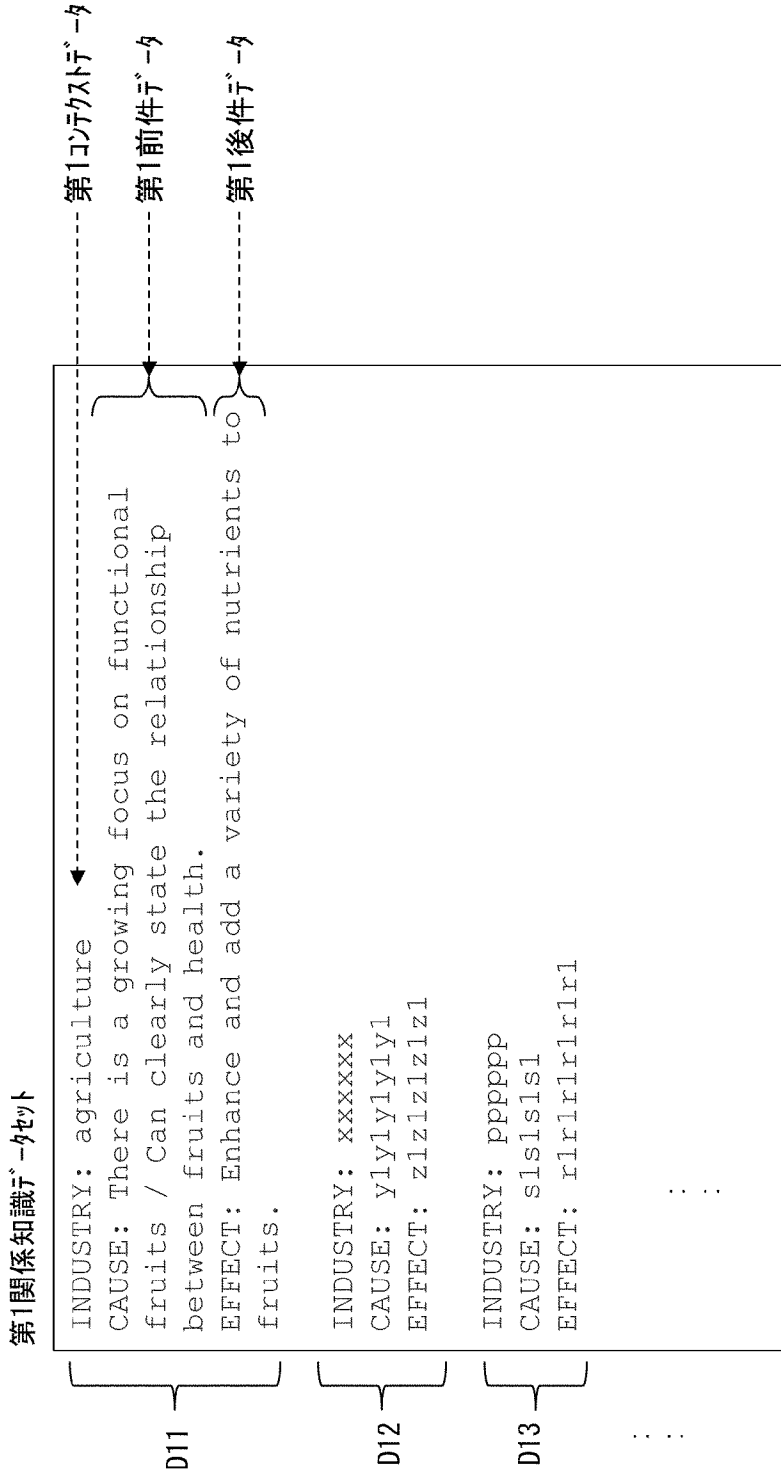


図 7

図 7



[図8]

図 8

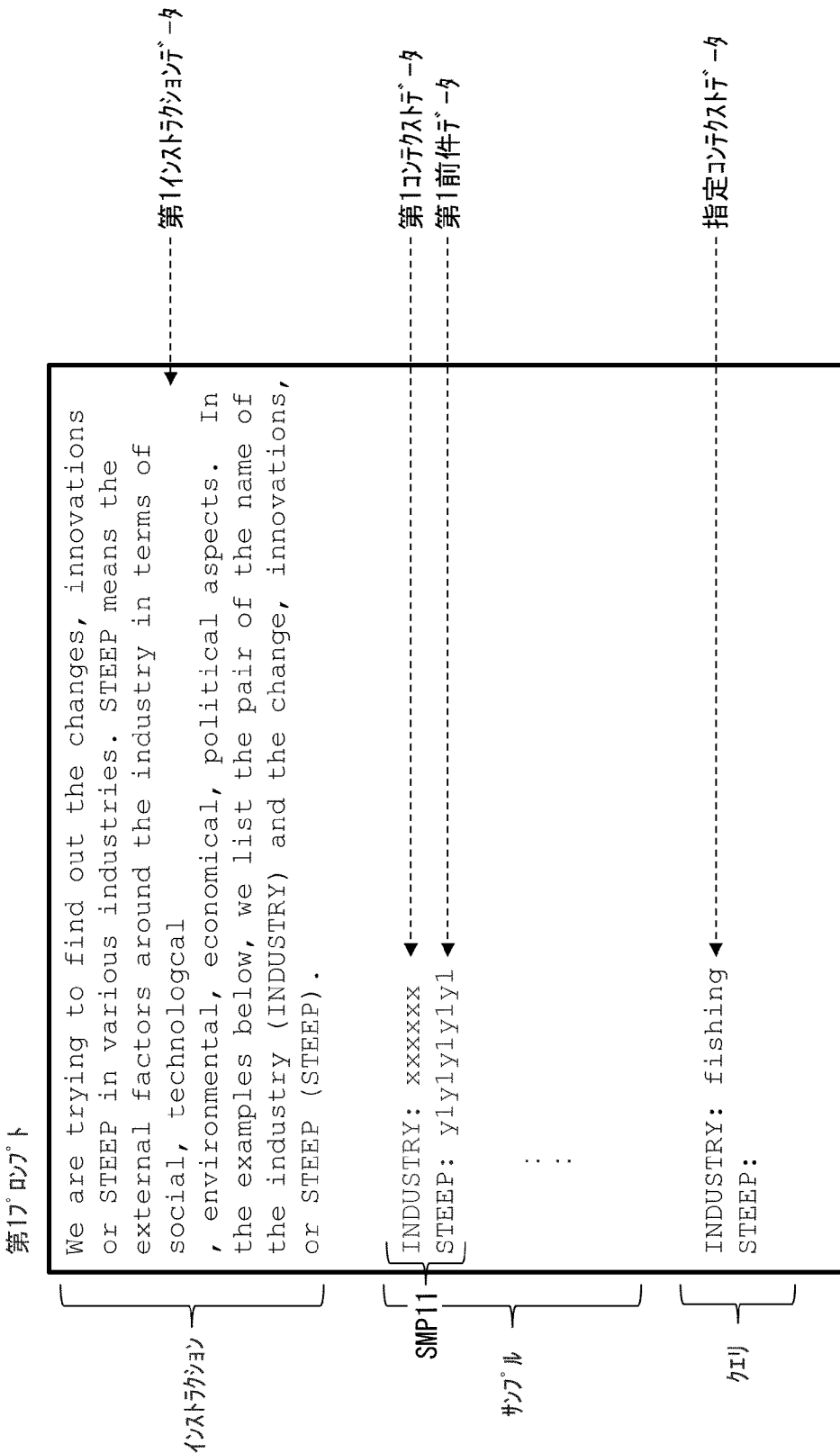
コンテキストリスト

```
agriculture  
fishing  
forestry
```

```
:  
:
```

[図9]

図9



[図10]

図 10

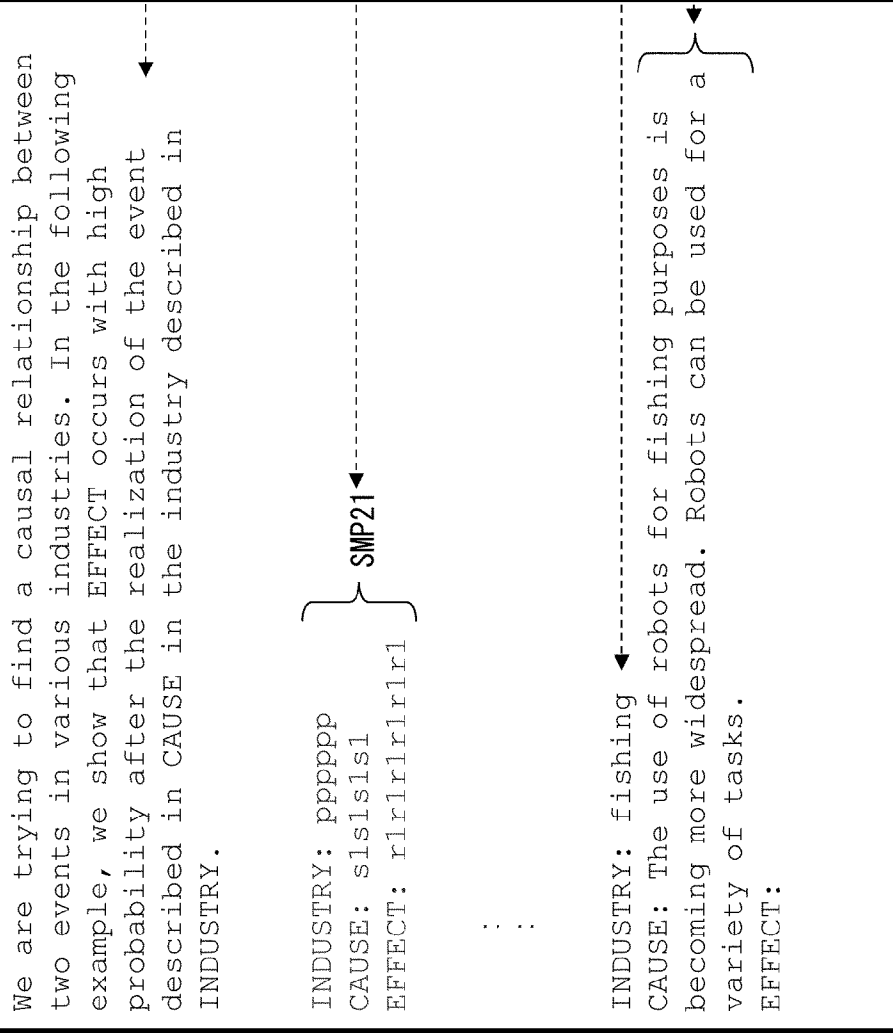
第2前件データリスト

第2コンテキストデータ	第2前件データ
fishing	The use of robots for fishing purposes is becoming more widespread. Robots can be used for a variety of tasks.
agriculture	x2x2x2x2 x2x2
:	:

[図11]

図 11

第2ポイント



インスタクション

サンプル

クイ

第2インスタクションデータ

第1関係知識データ

第2コンテキストデータ

第2前件データ

SMP21

[図12]

図 12

第2関係知識[°]-タセット

第2関係知識 [°] -タ		
第2コンテキスト [°] -タ	第2前件 [°] -タ	第2後件 [°] -タ
fishing	The use of robots for fishing purposes is becoming more widespread. Robots can be used for a variety of tasks.	The catches of fish will increase.
agriculture	x2x2x2x2 x2x2	y2y2y2 y2y2y2y2y
:	:	

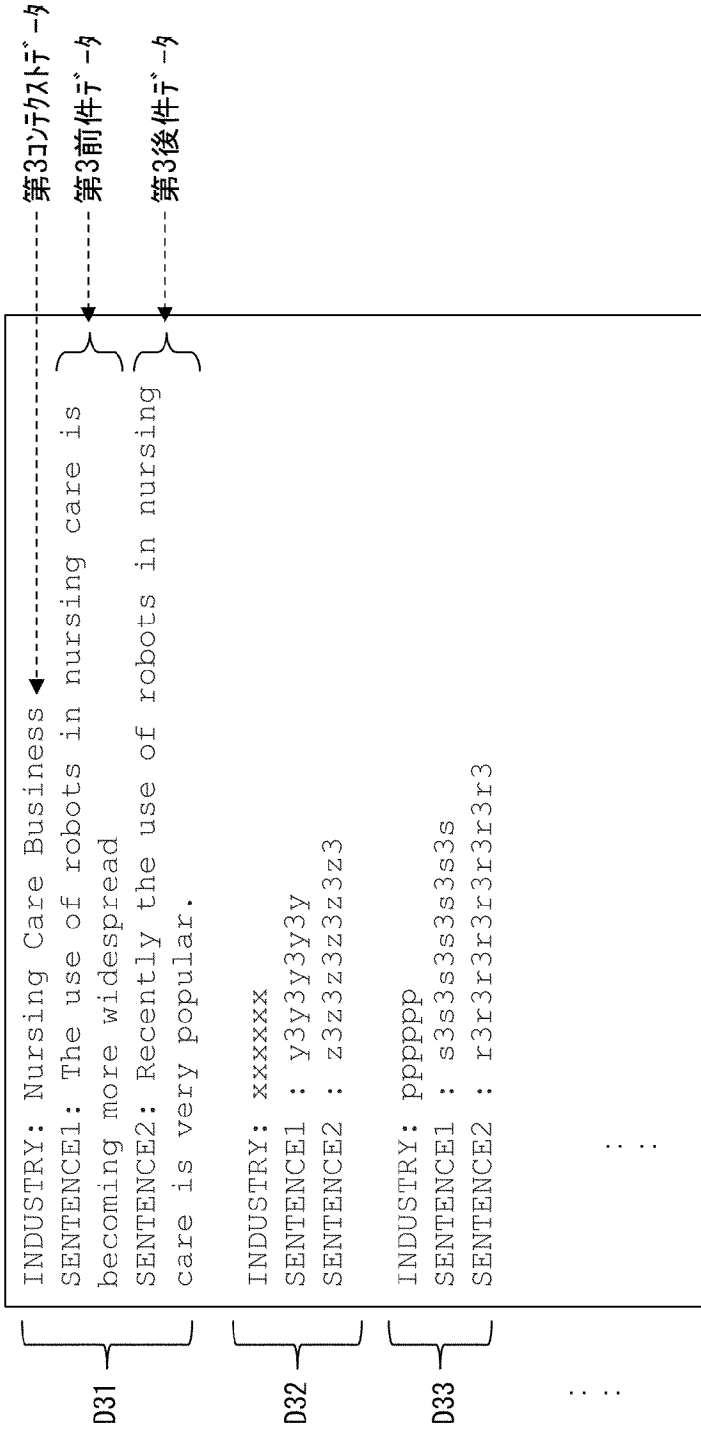
D21

D22

図 13

図13

第3関係知識データベース



[図14]

図 14

第37°ポイント

We are trying to determine if two sentences have the same meaning. In the following example, if "SENTENCE1" and "SENTENCE2" have the same meaning, then "LABEL" is "e". If "SENTENCE1" and "SENTENCE2" mean different things, then "LABEL" is "x".

INDUSTRY: Nursing Care Business
 SENTENCE1: The use of robots in nursing care is becoming more widespread
 SENTENCE2: Recently the use of robots in nursing care is very popular.
 Label: e

INDUSTRY: xxxxxxx
 SENTENCE1: Ylylylylyl
 SENTENCE2: zlz1z1z1z1
 Label: x

:
 :

INDUSTRY: fishing
 SENTENCE1: The use of robots for fishing purposes is becoming more widespread. Robots can be used for a variety of tasks.
 SENTENCE2: The catches of fish will increase.
 Label:

インスタレーション

サンプル

クエリ

第3インスタレーション

第3関係知識

第1関係知識

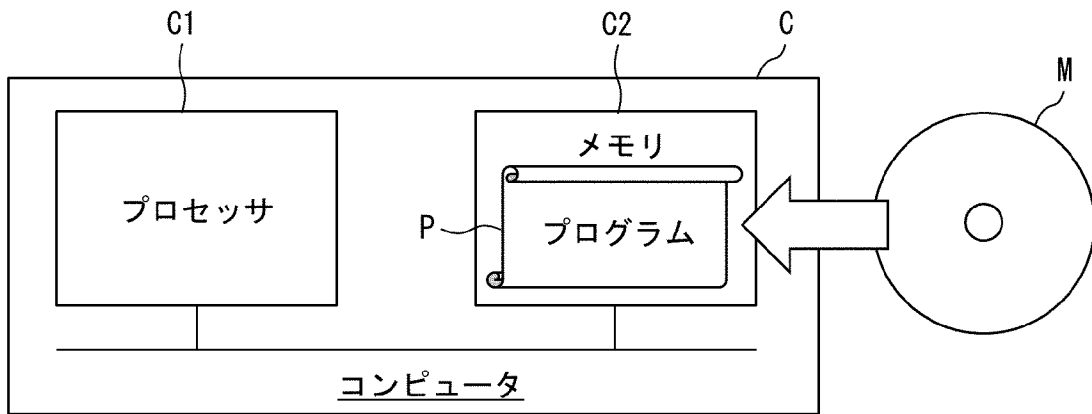
第2関係知識

SMP31

SMP32

[図15]

図 15



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2023/010286

A. CLASSIFICATION OF SUBJECT MATTER		
<i>G06N 5/022</i> (2023.01)i; <i>G06N 20/00</i> (2019.01)i FI: G06N5/022; G06N20/00		
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) G06N5/022; G06N20/00		
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 1971-2023 Registered utility model specifications of Japan 1996-2023 Published registered utility model applications of Japan 1994-2023		
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	WEST, Peter et al. Symbolic Knowledge Distillation: from General Language Models to Commonsense Models. arXiv [online]. November 2022, [retrieved on 29 May 2023], internet: <URL: https://arxiv.org/abs/2110.07178v2 >, 2110.07178v2 in particular, "3.2 Event Generation", "3.3 Inference Generation"	1-10
A	WO 2022/046087 A1 (SIEMENS AKTIENGESELLSCHAFT) 03 March 2022 (2022-03-03) paragraphs [0026]-[0031]	1-10
A	KR 10-2022-0067808 A (KT CORPORATION) 25 May 2022 (2022-05-25) paragraphs [0054]-[0055]	1-10
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“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)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&” document member of the same patent family</p>		
Date of the actual completion of the international search 30 May 2023		Date of mailing of the international search report 06 June 2023
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/JP2023/010286

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
WO 2022/046087 A1	03 March 2022	(Family: none)	
KR 10-2022-0067808 A	25 May 2022	(Family: none)	

A. 発明の属する分野の分類（国際特許分類（IPC）） G06N 5/022(2023.01)i; G06N 20/00(2019.01)i FI: G06N5/022; G06N20/00		
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）） G06N5/022; G06N20/00 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 - 2023年 日本国実用新案登録公報 1996 - 2023年 日本国登録実用新案公報 1994 - 2023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	WEST, Peter et al., Symbolic Knowledge Distillation: from General Language Models to Commonsense Models, arXiv [online], 2022.11, [検索日 2023.05.29], インターネット:<URL: https://arxiv.org/abs/2110.07178v2>, 2110.07178v2 特に、"3.2 Event Generation", "3.3 Inference Generation"	1-10
A	WO 2022/046087 A1 (SIEMENS AKTIENGESELLSCHAFT) 03.03.2022 (2022 - 03 - 03) [0026]-[0031]	1-10
A	KR 10-2022-0067808 A (KT CORPORATION) 25.05.2022 (2022 - 05 - 25) [0054]-[0055]	1-10
<input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input checked="" type="checkbox"/> 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 "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	国際調査報告の発送日	
30.05.2023	06.06.2023	
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） 坂庭 剛史 5B 1976 電話番号 03-3581-1101 内線 3545	

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/JP2023/010286

引用文献	公表日	パテントファミリー文献	公表日
WO 2022/046087 A1	03.03.2022	(ファミリーなし)	
KR 10-2022-0067808 A	25.05.2022	(ファミリーなし)	