

(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
2022年6月23日(23.06.2022)



(10) 国際公開番号

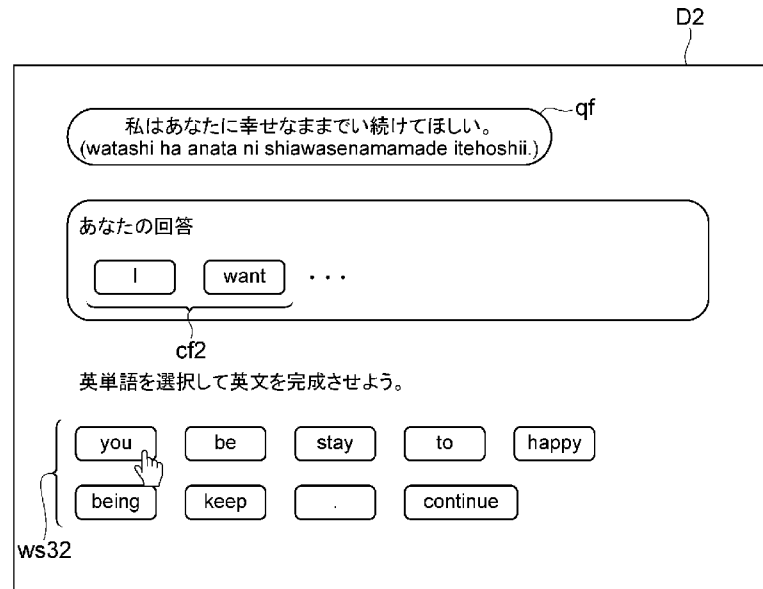
WO 2022/130940 A1

- (51) 国際特許分類:
G06F 40/44 (2020.01) *G06F 40/58* (2020.01)
G06F 40/47 (2020.01)
- (21) 国際出願番号: PCT/JP2021/043451
- (22) 国際出願日: 2021年11月26日(26.11.2021)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2020-207474 2020年12月15日(15.12.2020) JP
- (71) 出願人:株式会社 N T T ドコモ(NTT DOCOMO, INC.) [JP/JP]; 〒1006150 東京都千代田区永田町二丁目 1 1 番 1 号 Tokyo (JP).
- (72) 発明者: 村上 聡 一 朗 (MURAKAMI Soichiro); 〒1006150 東京都千代田区永田町二丁目 1 1 番 1 号 山王パークタワー 株式会社 N T T ドコモ 知的財産部内 Tokyo (JP).
- (74) 代理人: 長谷川 芳樹, 外 (HASEGAWA Yoshiki et al.); 〒1000005 東京都千代田区丸の内二丁目 1 番 1 号 丸の内 M Y P L A Z A (明治安田生命ビル) 9 階 創英国際特許法律事務所 Tokyo (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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN,

(54) Title: PRESENTATION DEVICE

(54) 発明の名称: 提示装置

[図9]



(57) Abstract: This presentation device is a device that presents candidates for words constituting a translated sentence that is a translation in a second language of an original sentence expressed in a first language. The presentation device comprises: a generation unit that generates a plurality of translated sentences in the second language on the basis of parallel-translation sentences output from a machine translation engine through input of original sentences, the machine translation engine being configured to output a parallel-translation sentence in the second language on the basis of an

[続葉有]

HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH,
KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, ME, MG, MK, MN, 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, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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, 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))

input sentence expressed in the first language; and a presentation unit that presents the words included in the plurality of translated sentences in an arrangement determined with a predetermined arrangement method.

(57) 要約：提示装置は、第1言語からなる原文の第2言語からなる翻訳文を構成する単語の候補を提示する装置であって、入力された第1言語の文に基づいて第2言語からなる対訳文を出力する機械翻訳エンジンに原文を入力することにより機械翻訳エンジンから出力される対訳文に基づいて、第2言語からなる複数の翻訳文を生成する生成部と、複数の翻訳文に含まれる単語を、所定の配列手法により配列して提示する提示部と、を備える。

明 細 書

発明の名称：提示装置

技術分野

[0001] 本発明は、第1言語からなる原文の第2言語からなる翻訳文を構成する単語の候補を提示する提示装置に関する。

背景技術

[0002] 第1言語（例えば日本語）で提示された原文の第2言語（例えば英語）の翻訳文を入力させることにより翻訳の学習をさせるシステムが知られている。このようなシステムにおいて、学習の障壁を低下させるために、翻訳文の自由入力に代えて、翻訳文を構成する単語を提示し、提示された単語に対する指定入力を受け付けることにより翻訳文を作成する方式が採用される場合がある。このような方式のシステムに適用しうる技術として、入力文を形態素解析により部分列に分割し、部分列に基づいて機械翻訳辞書から抽出された訳語を出力する装置が知られている（例えば、特許文献1参照）。

先行技術文献

特許文献

[0003] 特許文献1：特開2007-226729号公報

発明の概要

発明が解決しようとする課題

[0004] 従来における提示した単語に対する指定入力を受け付ける方式では、模範解答の翻訳文に用いられている単語しか提示できないという制約があり、多様な表現による解答を受け付けることができなかった。

[0005] そこで、本発明は、上記問題点に鑑みてなされたものであり、提示された単語に対する指定入力を受け付けることにより、第1言語の原文に対応する第2言語の翻訳文を作成するシステムにおいて、多様な表現による作文が可能となるような単語の候補を提示することを目的とする。

課題を解決するための手段

[0006] 上記課題を解決するために、本発明の一形態に係る提示装置は、第1言語からなる原文の第2言語からなる翻訳文を構成する単語の候補を提示する提示装置であって、入力された第1言語の文に基づいて第2言語からなる対訳文を出力する機械翻訳エンジンに原文を入力することにより機械翻訳エンジンから出力される対訳文に基づいて、第2言語からなる複数の翻訳文を生成する生成部と、複数の翻訳文に含まれる単語を、所定の配列手法により配列して提示する提示部と、を備える。

[0007] 上記の形態によれば、機械翻訳エンジンを用いて第1言語の原文に対応する第2言語の複数の翻訳文が生成され、生成された複数の翻訳文に含まれる単語が提示される。従って、模範解答を構成しうるような一の特定の翻訳文に含まれる単語以外の単語を、単語の候補として提示できる。このように提示された単語の候補に対する指定入力を受け付けることにより、多様な表現による作文を行うことが可能となる。

発明の効果

[0008] 提示された単語に対する指定入力を受け付けることにより、第1言語の原文に対応する第2言語の翻訳文を作成するシステムにおいて、多様な表現による作文が可能となるような単語の候補を提示することが可能となる。

図面の簡単な説明

[0009] [図1]本実施形態の提示装置の機能的構成を示すブロック図である。

[図2]提示装置及び文生成装置のハードブロック図である。

[図3]原文記憶部に記憶されている原文の例を示す図である。

[図4]本実施形態の提示装置における原文に基づく単語の候補の提示に至る処理の概略を示す図である。

[図5]機械翻訳エンジンの一例であって、エンコーダデコーダモデルを含むニューラル機械翻訳エンジンの例を示す図である。

[図6]既知のビームサーチの手法の概念を概略的に示す図である。

[図7]エンコーダデコーダモデルを含むニューラル機械翻訳エンジンにより、提示される単語の候補が指定入力された単語に応じて動的に変更されて出力

される処理の概略を示す図である。

[図8]単語の候補の提示に応じて受け付けられた指定入力に基づいて生成された作成文を提示する表示画面の例を示す図である。

[図9]単語の候補の提示に応じて受け付けられた指定入力に基づいて生成された作成文を提示する表示画面の例を示す図である。

[図10]単語の候補の提示に応じて受け付けられた指定入力に基づいて生成された作成文を提示する表示画面の例を示す図である。

[図11]提示装置における提示方法の処理内容を示すフローチャートである。

[図12]提示装置における提示方法の処理内容を示すフローチャートである。

[図13]提示プログラムの構成を示す図である。

発明を実施するための形態

[0010] 本発明に係る提示装置の実施形態について図面を参照して説明する。なお、可能な場合には、同一の部分には同一の符号を付して、重複する説明を省略する。

[0011] 図1は、本実施形態に係る提示装置を含むシステムの装置構成及び提示装置の機能的構成を示す図である。提示装置10は、第1言語からなる原文の第2言語からなる翻訳文を構成する単語の候補を提示する装置である。

[0012] 提示装置10は、原文記憶部20といった記憶手段にアクセス可能に構成されている。原文記憶部20は、提示装置10内に構成されてもよいし、図1に示されるように、提示装置10の外部に、提示装置からアクセス可能な別の装置として構成されてもよい。

[0013] 提示装置10は、端末30とネットワークを介して通信可能に構成されている。端末30は、ユーザにより操作される装置である。端末30は、提示装置10から送信された情報の表示、ユーザによる入力の受け付け及び入力された情報の提示装置10への送信等を実施する。

[0014] なお、図1に示す例では、提示装置10と端末30とは別の装置として示されているが、提示装置10の全ての機能部が端末30に構成されることにより端末30が提示装置10として機能することとしてもよいし、提示装置

１０の機能部のうちの一部の機能部が端末３０に構成されてもよい。

[0015] 提示装置１０は、機能的には、取得部１１、生成部１２、提示部１３、指定入力受付部１４、問題文提示部１５及び作文部１６を備える。これらの各機能部１１～１６は、一つの装置（コンピュータ）に構成されてもよいし、複数の装置に分散されて構成されてもよい。

[0016] なお、図１に示したブロック図は、機能単位のブロックを示している。これらの機能ブロック（構成部）は、ハードウェア及びソフトウェアの少なくとも一方の任意の組み合わせによって実現される。また、各機能ブロックの実現方法は特に限定されない。すなわち、各機能ブロックは、物理的又は論理的に結合した１つの装置を用いて実現されてもよいし、物理的又は論理的に分離した２つ以上の装置を直接的又は間接的に（例えば、有線、無線などを用いて）接続し、これら複数の装置を用いて実現されてもよい。機能ブロックは、上記１つの装置又は上記複数の装置にソフトウェアを組み合わせることで実現されてもよい。

[0017] 機能には、判断、決定、判定、計算、算出、処理、導出、調査、探索、確認、受信、送信、出力、アクセス、解決、選択、選定、確立、比較、想定、期待、見做し、報知（broadcasting）、通知（notifying）、通信（communicating）、転送（forwarding）、構成（configuring）、再構成（reconfiguring）、割り当て（allocating、mapping）、割り振り（assigning）などがあるが、これらに限られない。たとえば、送信を機能させる機能ブロック（構成部）は、送信部（transmitting unit）や送信機（transmitter）と称される。いずれも、上述したとおり、実現方法は特に限定されない。

[0018] 例えば、本発明の一実施の形態における提示装置１０は、コンピュータとして機能してもよい。図２は、本実施形態に係る提示装置１０のハードウェア構成の一例を示す図である。提示装置１０は、物理的には、プロセッサ１００１、メモリ１００２、ストレージ１００３、通信装置１００４、入力装置１００５、出力装置１００６、バス１００７などを含むコンピュータ装置として構成されてもよい。

- [0019] なお、以下の説明では、「装置」という文言は、回路、デバイス、ユニットなどに読み替えることができる。提示装置 10 のハードウェア構成は、図に示した各装置を 1 つ又は複数含むように構成されてもよいし、一部の装置を含まずに構成されてもよい。
- [0020] 提示装置 10 における各機能は、プロセッサ 1001、メモリ 1002 などのハードウェア上に所定のソフトウェア（プログラム）を読み込ませることで、プロセッサ 1001 が演算を行い、通信装置 1004 による通信や、メモリ 1002 及びストレージ 1003 におけるデータの読み出し及び／又は書き込みを制御することで実現される。
- [0021] プロセッサ 1001 は、例えば、オペレーティングシステムを動作させてコンピュータ全体を制御する。プロセッサ 1001 は、周辺装置とのインターフェース、制御装置、演算装置、レジスタなどを含む中央処理装置（CPU: Central Processing Unit）で構成されてもよい。例えば、図 1 に示した各機能部 11～16 などは、プロセッサ 1001 で実現されてもよい。
- [0022] また、プロセッサ 1001 は、プログラム（プログラムコード）、ソフトウェアモジュールやデータを、ストレージ 1003 及び／又は通信装置 1004 からメモリ 1002 に読み出し、これらに従って各種の処理を実行する。プログラムとしては、上述の実施の形態で説明した動作の少なくとも一部をコンピュータに実行させるプログラムが用いられる。例えば、提示装置 10 の各機能部 11～16 は、メモリ 1002 に格納され、プロセッサ 1001 で動作する制御プログラムによって実現されてもよい。上述の各種処理は、1 つのプロセッサ 1001 で実行される旨を説明してきたが、2 以上のプロセッサ 1001 により同時又は逐次に実行されてもよい。プロセッサ 1001 は、1 以上のチップで実装されてもよい。なお、プログラムは、電気通信回線を介してネットワークから送信されてもよい。
- [0023] メモリ 1002 は、コンピュータ読み取り可能な記録媒体であり、例えば、ROM (Read Only Memory)、EPROM (Erasable Programmable ROM)、EEPROM (Electrically Erasable Programmable ROM)、

R A M (Random Access Memory) などの少なくとも1つで構成されてもよい。メモリ1002は、レジスタ、キャッシュ、メインメモリ（主記憶装置）などと呼ばれてもよい。メモリ1002は、本発明の一実施の形態に係る提示方法を実施するために実行可能なプログラム（プログラムコード）、ソフトウェアモジュールなどを保存することができる。

[0024] ストレージ1003は、コンピュータ読み取り可能な記録媒体であり、例えば、C D - R O M (Compact Disc ROM) などの光ディスク、ハードディスクドライブ、フレキシブルディスク、光磁気ディスク(例えば、コンパクトディスク、デジタル多用途ディスク、B l u - r a y (登録商標) ディスク)、スマートカード、フラッシュメモリ(例えば、カード、スティック、キードライブ)、フロッピー（登録商標）ディスク、磁気ストリップなどの少なくとも1つで構成されてもよい。ストレージ1003は、補助記憶装置と呼ばれてもよい。上述の記憶媒体は、例えば、メモリ1002及び／又はストレージ1003を含むデータベース、サーバその他の適切な媒体であってもよい。

[0025] 通信装置1004は、有線及び／又は無線ネットワークを介してコンピュータ間の通信を行うためのハードウェア（送受信デバイス）であり、例えばネットワークデバイス、ネットワークコントローラ、ネットワークカード、通信モジュールなどともいう。

[0026] 入力装置1005は、外部からの入力を受け付ける入力デバイス（例えば、キーボード、マウス、マイクロフォン、スイッチ、ボタン、センサなど）である。出力装置1006は、外部への出力を実施する出力デバイス（例えば、ディスプレイ、スピーカー、LEDランプなど）である。なお、入力装置1005及び出力装置1006は、一体となった構成（例えば、タッチパネル）であってもよい。

[0027] また、プロセッサ1001やメモリ1002などの各装置は、情報を通信するためのバス1007で接続される。バス1007は、単一のバスで構成されてもよいし、装置間で異なるバスで構成されてもよい。

[0028] また、提示装置 10 は、マイクロプロセッサ、デジタル信号プロセッサ（DSP: Digital Signal Processor）、ASIC (Application Specific Integrated Circuit)、PLD (Programmable Logic Device)、FPGA (Field Programmable Gate Array) などのハードウェアを含んで構成されてもよく、当該ハードウェアにより、各機能ブロックの一部又は全てが実現されてもよい。例えば、プロセッサ 1001 は、これらのハードウェアの少なくとも 1 つで実装されてもよい。

[0029] 再び図 1 を参照して、提示装置 10 の機能部を説明する。取得部 11 は、原文を原文記憶部 20 から取得する。原文は、第 1 言語からなる文であって、提示装置 10 が和英翻訳に適用される場合には、日本語の文である。

[0030] 図 3 は、原文記憶部 20 に記憶されている原文の例を示す図である。原文記憶部 20 は、各原文を識別する ID に関連付けられた原文を複数記憶しており、例えば、「彼女は牛乳をいくらか欲しがっています。（kanojo ha gyunyuu wo ikuraka hoshigatteimasu.）」、「私は公園の中を走っています。（watashi ha koen no naka wo hashitteimasu.）」等の第 1 言語の文を記憶している。

[0031] 生成部 12 は、機械翻訳エンジンに原文を入力することにより機械翻訳エンジンから出力される対訳文に基づいて、第 2 言語からなる複数の翻訳文を生成する。機械翻訳エンジンは、入力された第 1 言語の文に基づいて第 2 言語からなる対訳文を出力する。

[0032] 提示部 13 は、複数の翻訳文に含まれる単語を、所定の配列手法により配列して提示する。具体的には、提示部 13 は、複数の翻訳文に含まれる単語を、端末 30 のディスプレイに表示させる。

[0033] 図 4 は、提示装置 10 における原文に基づく単語の候補の提示に至る処理の概略を示す図である。図 4 に示されるように、生成部 12 は、原文 s f 0 「彼女は牛乳をいくらか欲しがっています。（kanojo ha gyunyuu wo ikuraka hoshigatteimasu.）」を機械翻訳エンジン m t 0 に入力する。生成部 12 は、機械翻訳エンジン m t 0 から出力される対訳文に基づいて、複数の翻訳文

t f を生成する。なお、図 4 に示される例では、機械翻訳エンジン m t 0 は、ニューラルネットワークを含んで構成されるニューラル機械翻訳エンジンであって、ビームサーチの手法により、対訳としての尤もらしさを示す尤度（第 2 の確率）と共に対訳を出力する。生成部 1 2 は、例えば、尤度の上位 4 位までの対訳文に基づいて、複数の翻訳文 t f を生成する。

[0034] 提示部 1 3 は、複数の翻訳文 t f に含まれる単語 w s を、所定の配列手法により配列して提示する。図 4 に示されるように、提示部 1 3 は、複数の翻訳文 t f に含まれる 4 つの翻訳文に含まれる単語 w s（needs／．／She／some／get／to／wants／is／milk／wanting）を端末 3 0 のディスプレイに表示させる。

[0035] 提示部 1 3 は、図 4 に示されるように、複数の翻訳文 t f に含まれる単語をランダムに配列して提示してもよい。端末 3 0 において複数の翻訳文 t f に含まれる単語をランダムに配列して提示させることにより、提示された単語の配列順に基づいて、翻訳文における単語の配列順を認識することを不可能にさせることができる。

[0036] 従来における、翻訳文を構成する単語を提示し、提示された単語に対する指定入力を受け付けることにより翻訳文を作成する技術では、単語 w 0 に示されるように、原文 s f 0 の模範解答の翻訳文（尤度が最大の翻訳文）（She wants some milk.）に用いられている単語しか提示できなかった。これに対して、本実施形態の提示装置 1 0 では、単語 w s に示されるように、複数の翻訳文 t f に含まれる単語が提示されるので、多様な表現による作文が可能となるような単語の候補を提示できる。

[0037] 本実施形態の機械翻訳エンジンは、学習済みの言語モデルを含むニューラル機械翻訳エンジンであってもよい。言語モデルは、ニューラルネットワークを含み、第 1 言語の文の入力に基づいて、対訳文を構成する単語列の各単語に対応するタイムステップごとに 1 以上の単語を、前記対訳文に用いられる単語としての確からしさを示す確率と共に順次出力する。生成部 1 2 は、順次出力された単語の単語列からなる対訳文を取得する。

- [0038] 機械翻訳エンジンは、言語モデルを含むエンコーダデコーダモデルにより構成されてもよい。図5は、機械翻訳エンジンの一例であって、エンコーダデコーダモデルを含むニューラル機械翻訳エンジンの例を示す図である。
- [0039] 機械翻訳エンジン $m t 1$ は、エンコーダ $e n 1$ 及びデコーダ $d e 1$ からなるエンコーダデコーダモデルを含む。エンコーダデコーダモデルを構成するニューラルネットワークは限定されないが、例えばリカレントニューラルネットワーク（RNN: Recurrent Neural Network）、または、トランスフォーマ（Transformer）と言われるニューラルネットワークであってもよい。機械翻訳エンジン $m t 1$ を構成するニューラルネットワークは、第1言語の文と第2言語の対訳とのペアからなる学習データを用いた機械学習により学習済みである。
- [0040] 生成部12は、第1言語の原文 $s f 1$ をエンコーダ $e n 1$ に入力する。具体的には、生成部12は、形態素解析等により原文 $s f 1$ を単語に分割し、分割された単語を対応する単語ベクトルに変換（Embedding）し、単語ベクトルをエンコーダ $e n 1$ に順次入力する。エンコーダ $e n 1$ は、第1データ a に基づく計算結果を示すベクトル（例えば、中間層の出力、ソースターゲットアテンション等）をデコーダ $d e 1$ に出力する。
- [0041] デコーダ $d e 1$ は、エンコーダ $e n 1$ からのベクトル及び出力の開始を示す所定の開始記号 $s s$ （ベクトル）の入力に基づいて、各タイムステップ $t s$ において単語群 $w t$ の系列 $w t 1, w t 2, w t 3, \dots, w t n$ を順次出力する。さらに具体的には、デコーダ $d e 1$ は、各タイムステップ $t s$ において、複数の単語 $c w$ を、対訳文に用いられる単語としての確からしさ（尤もらしさ）を示す確率 $w p$ と共に出力する。デコーダ $d e 1$ から出力の終了を意味する終端記号 $s e$ が出力されたら、それまでに各タイムステップにおいて順次出力された単語の系列により対訳が構成される。
- [0042] 従来における、一意の対訳を得るための翻訳エンジンでは、各タイムステップにおいて最も確率 $w p$ が高い単語 $c w$ を、次のタイムステップにおけるデコーダ $d e 1$ の入力とする。例えば、タイムステップ $t s 1$ において出力

された単語群 w_t において、単語 c_{w11} が最も高い確率 w_{p11} を有するので、タイムステップ t_{s2} では、単語 c_{w11} がデコーダ d_e に入力される。

[0043] これに対して、本実施形態では、ビームサーチ及び $Top-K$ サンプリング等の既知の手法により、あるタイムステップ t_s において出力された単語群 w_t から、対訳文を構成する単語としての確からしさを示す第1の確率に基づいて、複数の単語が抽出及び保持され、保持された複数の単語のそれぞれを次のタイムステップ t_s におけるデコーダ d_e への入力として、複数の単語のそれぞれの入力に基づく単語群 w_t が得られる。

[0044] 図6は、既知のビームサーチの手法の概念を概略的に示す図である。ビームサーチは、尤もらしい単語の組み合わせを対訳として抽出するための既知の探索アルゴリズムである。図6のビームサーチの例において、探索の幅を規定するビーム幅 b_w は3であり、丸は単語を表し、丸の中の数字は単語の尤度のスコア（第1の確率 w_p ）を表す。

[0045] 生成部12は、デコーダ d_e のタイムステップ t_{s1} において、出力された単語群 w_t のうち、スコアの上位3つの単語 c_w を抽出及び保持する。タイムステップ t_{s2} において、生成部12は、タイムステップ t_{s1} において保持した単語 c_{w1} のそれぞれをデコーダ d_e に入力し、単語群 w_t を取得する。生成部12は、それまでのタイムステップにおいて抽出及び保持された単語からなる単語列の尤度のスコア（第2の確率）の上位3位までを構成する単語を抽出及び保持する。引き続くタイムステップ $t_{s3} \sim t_{s5}$ において、生成部12は、前タイムステップにおいて保持した単語のそれぞれをデコーダ d_e に入力し、同様に単語を抽出及び保持する。

[0046] 生成部12は、各タイムステップにおいて、単語列の尤度のスコアに基づいて、ビーム幅 b_w として規定される数の単語をスコアの上位から抽出及び保持し、前タイムステップにおいて保持された単語に基づいてデコーダ d_e から出力された単語の抽出及び保持を繰り返し、各単語列の探索において終端記号 s_e が出力されるまで、当該単語列のための単語を抽出する。

[0047] 生成部12は、ビームサーチの手法をエンコーダデコーダモデルに適用して得られた、単語列の尤度のスコア（単語列の対訳文としての確からしさを示す第2の確率）が上位所定数の単語列からなる対訳文を、複数の翻訳文として出力する。このように、文頭から順次抽出された単語により構成される単語列に関する第2の確率が算出され、第2の確率の上位所定数の対訳文が出力されるので、原文に対応する翻訳文としての尤度が高い単語列を翻訳文として出力できる。また、出力された複数の翻訳文に含まれる単語が提示されるので、多様な表現の翻訳文に対応した単語の候補の提示が可能となる。

[0048] また、生成部12は、既知のTop-Kサンプリングの手法をエンコーダデコーダモデルに適用して得られる単語列からなる対訳文を、翻訳文として生成及び出力してもよい。即ち、生成部12は、あるタイムステップ t_s において、対訳文を構成する単語としての確からしさを示す確率 w_p （尤度、第1の確率）の上位所定数の単語の中からそれらの確率に基づき重み付きランダムサンプリング（重み付き無作為抽出）を行い、抽出した単語を保持する。次のタイムステップ t_s において、生成部12は、前のタイムステップ t_s において保持された単語をデコーダ d_e1 に入力し、デコーダ d_e1 から出力された単語群 w_t のうちの、第1の確率の上位所定数の単語を抽出及び保持する。

[0049] 各タイムステップにおいて、生成部12は、第1の確率の上位所定数の単語を順次抽出し、抽出された単語の単語列により構成される複数の対訳文を、複数の翻訳文として出力する。従って、原文に対応する翻訳文としての尤度が高い単語列を翻訳文として出力できる。また、各タイムステップにおいて抽出する単語の所定数を調整する（多くする）ことにより、尤度が低い翻訳文を得ることが可能となる、そのような翻訳文に含まれる単語が提示されることにより、誤りを含んだ単語の候補の提示が可能となる。

[0050] 再び図1を参照して、指定入力受付部14は、提示部13により提示された複数の単語のうちの少なくとも一つの単語を指定する指定入力を受け付ける。生成部12は、指定入力受付部14により受け付けられた単語である指

定単語をエンコーダデコーダモデルのデコーダに入力し、指定単語の入力に引き続いてデコーダの各タイムステップにおいて第1の確率に基づいて複数出力される単語を順次抽出し、順次抽出された単語からなる単語列に基づいて、複数の翻訳文を生成する。

[0051] 図7は、エンコーダデコーダモデルを含むニューラル機械翻訳エンジンにより、提示される単語の候補が指定入力された単語に応じて動的に変更されて出力される処理の概略を示す図である。

[0052] 図7に示される例では、生成部12は、原文s f 2（彼はサッカーの試合でけがをした。（kare ha sakka no shiai de kega wo shita.））をエンコーダデコーダモデルからなる機械翻訳エンジンm t 2のエンコーダe n 2に入力することにより、機械翻訳エンジンm t 2から出力される対訳文に基づいて、複数の翻訳文を生成する。提示部13は、生成部12により生成された複数の翻訳文に含まれる単語を、例えば端末30のディスプレイに表示させる。

[0053] 端末30において、表示された複数の単語のうちの、単語「H e」及びそれに引き続く単語「g o t」がユーザにより指定されると、指定入力受付部14は、単語「H e」、「g o t」を指定する指定入力を受け付ける。

[0054] 生成部12は、指定入力受付部14により受け付けられた単語である指定単語i w（「H e」、「g o t」）を、当該指定単語i wに対応する所定のタイムステップt sにおいて、前タイムステップt sにおいて出力された単語に代えてデコーダd e 2に入力する。

[0055] 生成部12は、図5等を参照して説明したように、指定単語i wの入力に引き続いてデコーダd e 2の各タイムステップt sにおいて第1の確率（対訳文を構成する単語としての尤もらしさ）に基づいて複数出力される単語を順次抽出し、順次抽出された単語からなる単語列に基づいて、複数の翻訳文t f 2を生成する。図7に示される例においては、翻訳文t f 2に含まれる各翻訳文は、原文s f 2の対訳文における指定単語i w（「H e」、「g o t」）に引き続く文の部分である。

- [0056] 提示部 1 3 は、指定入力受付部 1 4 により指定単語 $i w$ が受け付けられるたびに、生成部 1 2 により生成された複数の翻訳文に含まれる単語を提示する。図 7 に示される例においては、提示部 1 3 は、翻訳文 $t f 2$ に含まれる各翻訳文に含まれる単語 $w s 2$ を、例えば端末 3 0 のディスプレイに表示させる。
- [0057] また、機械翻訳エンジン $m t 0$ は、既知の統計的機械翻訳により原文 $s f 0$ に対応する対訳文を出力する統計的機械翻訳エンジンであってもよい。統計的機械翻訳エンジンは、翻訳モデルと言語モデルを含む。
- [0058] 統計的機械翻訳エンジンにおける翻訳モデルは、第 1 言語の単語及びフレーズと、第 2 言語の対訳とからなるコーパスにより機械学習された学習済みのモデルであって、第 1 言語の文を構成する単語の第 2 言語による対訳の候補である単語対訳候補を、対訳としての確からしさを示す第 3 の確率と共に出力する。
- [0059] 統計的機械翻訳エンジンにおける言語モデルは、第 2 言語の文からなるコーパスにより文脈を機械学習された学習済みのモデルであって、翻訳モデルから出力された単語対訳候補に基づいて、第 2 言語の文章として自然な単語を選択しながら対訳文を生成する。言語モデルは、生成された対訳文において選択された各単語の対訳文を構成する単語としての尤度（第 4 の確率）を、各単語と併せて出力する。
- [0060] 生成部 1 2 は、原文 $s f 0$ を統計的機械翻訳エンジンからなる機械翻訳エンジン $m t 0$ に入力し、第 3 の確率及び第 4 の確率に基づいて、複数の翻訳文を生成する。従来の一意的翻訳文を生成するタスクにおいては、第 3 及び第 4 の確率において最高のスコアを有する単語及び単語の配列を用いることにより一意的翻訳文を得ることができる。これに対して本実施形態では、生成部 1 2 は、翻訳モデル及び言語モデルのそれぞれの探索の過程において、第 3 及び第 4 の確率のスコアが最上位の単語等だけではなく、上位所定数のスコアを有する単語等を採用することにより、複数の翻訳文を生成できる。また、生成部 1 2 は、翻訳モデル及び言語モデルのそれぞれの探索の過程に

において、スコアに基づいて抽出された単語等を意図的に他の単語等（予め設定された類似単語等）に置換することにより、複数の翻訳文を生成することとしてもよい。

[0061] 再び図 1 を参照して、問題文提示部 1 5 は、第 1 言語の原文により構成される問題文を提示する。

[0062] 作文部 1 6 は、提示部 1 3 により提示された単語に対する指定入力を受け付け、指定入力された単語を指定入力の順に応じて配列してなる作成文を生成し、作成文を提示する。

[0063] 図 8 ～図 1 0 を参照して、問題文の提示及び作成文の生成及び提示について説明する。図 8 ～図 1 0 は、単語の候補の提示に応じて受け付けられた指定入力に基づいて生成された作成文を提示する表示画面の例を示す図である。

[0064] 図 8 の表示画面 D 1 に示されるように、問題文提示部 1 5 は、第 1 言語の原文により構成される問題文 q f（私はあなたに幸せなままでい続けてほしい。（*watashi ha anata ni shiawasenamamade itehoshii.*））を提示する。具体的には、問題文提示部 1 5 は、端末 3 0 のディスプレイにおける表示画面 D 1 に問題文 q f を表示させる。提示部 1 3 は、問題文 q f として提示された原文に基づいて生成部 1 2 により生成された複数の翻訳文に含まれる単語 w s 3 1 を表示画面 D 1 に表示させる。

[0065] 指定入力受付部 1 4 は、提示された単語 w s 3 1 に対する指定入力を受け付けることができる。図 8 に示される例では、指定入力受付部 1 4 は、単語 w s 3 1 のうちの単語「I」に対する指定入力を受け付けることができる。

[0066] 図 9 に示される表示画面 D 2 は、指定入力受付部 1 4 が、単語「I」に対する指定入力に続いて単語「want」に対する指定入力を受け付けた場合の例である。図 9 に示されるように、作文部 1 6 は、指定入力された単語「I」，「want」を指定入力の順に配列してなる作成部 c f 2 を生成し、表示画面 D 2 おいて表示させる。また、表示画面 D 2 において、提示部 1 3 は、単語「I」，「want」の指定入力に応じて生成部 1 2 により再生成

された複数の翻訳文に含まれる単語w s 3 2を提示する。

[0067] 図10に示される表示画面D3は、指定入力受付部14が、単語「I」, 「want」, 「you」, 「to」, 「keep」に対する指定入力を受け付けた場合の例である。図10に示されるように、作文部16は、指定入力された単語「I」, 「want」, 「you」, 「to」, 「keep」を指定入力の順に配列してなる作成部cf3を生成し、表示画面D3において表示させる。

[0068] また、表示画面D3において、提示部13は、単語「I」, 「want」, 「you」, 「to」, 「keep」の指定入力に応じて生成部12により再生成された複数の翻訳文に含まれる単語w s 3 3を提示する。特に、単語w s 3 4に示されるように、指定入力された単語からなる単語列を制約として、単語列に引き続く尤もらしい翻訳文が再生成されて、再生成された翻訳文に基づいて単語が提示されることにより、単語の指定入力に応じて表示される単語を動的に変更できる。

[0069] 図11は、提示装置10における提示方法の処理内容を示すフローチャートである。

[0070] ステップS1において、取得部11は、原文を例えば原文記憶部20から取得する。

[0071] ステップS2において、生成部12は、機械翻訳エンジンに原文を入力することにより機械翻訳エンジンから出力される対訳文に基づいて、第2言語からなる複数の翻訳文を生成する。

[0072] ステップS3において、提示部13は、複数の翻訳文に含まれる単語を、所定の配列手法により配列して提示する。具体的には、提示部13は、複数の翻訳文に含まれる単語を、端末30のディスプレイに表示させる。

[0073] ステップS4において、指定入力受付部14は、提示部13により提示された複数の単語のうちの少なくとも一つの単語を指定する指定入力を受け付ける。

[0074] ステップS5において、指定入力された単語を指定入力の順に応じて配列

してなる作成文を生成し、生成した作成文を提示する。

[0075] ステップS 6において、提示装置10は、文の作成が終了したか否かを判断する。文の作成の終了は、例えば、ユーザによる入力に基づいて判断されてもよいし、所定の指定入力（例えばピリオド）があったことに基づいて判断されてもよい。文の作成が終了したと判断された場合には、処理は終了する。一方、文の作成が終了したと判断されなかった場合には、処理はステップS 4に戻る。

[0076] 図12は、提示装置における提示方法の処理内容の他の例を示すフローチャートである。図12のフローチャートにおけるステップS 11～ステップS 15の処理は、図11のフローチャートにおけるステップS 1～ステップS 5の処理と同様であるので、説明を省略する。

[0077] ステップS 16において、提示装置10は、文の作成が終了したか否かを判断する。文の作成が終了したと判断された場合には、処理は終了する。一方、文の作成が終了したと判断されなかった場合には、処理はステップS 17に進む。

[0078] ステップS 17において、生成部12は、ステップS 14において受け付けられた指定単語を、当該指定単語に対応する所定のタイムステップにおいて、前タイムステップt sにおいて出力された単語に代えてデコーダに入力する。そして、生成部12は、指定単語の入力に引き続いてデコーダの各タイムステップにおいて出力される単語を順次取得し、順次取得された単語からなる単語列に基づいて、複数の翻訳文を再生成する。

[0079] ステップS 18において、提示部13は、生成部12により再生成された複数の翻訳文に含まれる単語を提示する。そして、処理はステップS 14に戻る。

[0080] 次に、図13を参照して、コンピュータを、本実施形態の提示装置10として機能させるための提示プログラムについて説明する。

[0081] 図13は、提示プログラムの構成を示す図である。提示プログラムP 1は、提示装置10における提示処理を統括的に制御するメインモジュールm 1

0、取得モジュールm 1 1、生成モジュールm 1 2、提示モジュールm 1 3、指定入力受付モジュールm 1 4、問題文提示モジュールm 1 5及び作文モジュールm 1 6を備えて構成される。そして、各モジュールm 1 1～m 1 6により、取得部1 1、生成部1 2、提示部1 3、指定入力受付部1 4、問題文提示部1 5及び作文部1 6のための各機能が実現される。

[0082] なお、提示プログラムP 1は、通信回線等の伝送媒体を介して伝送される態様であってもよい、図1 3に示されるように、記録媒体M 1に記憶される態様であってもよい。

[0083] 以上説明した本実施形態の提示装置1 0、提示方法及び提示プログラムP 1によれば、機械翻訳エンジンを用いて第1言語の原文に対応する第2言語の複数の翻訳文が生成され、生成された複数の翻訳文に含まれる単語が提示される。従って、模範解答を構成しうるような一の特定の翻訳文に含まれる単語以外の単語を、単語の候補として提示できる。このように提示された単語の候補に対する指定入力を受け付けることにより、多様な表現による作文を行うことが可能となる。

[0084] また、別の形態に係る提示装置では、機械翻訳エンジンは、ニューラルネットワークを含み、第1言語の文の入力に基づいて、対訳文を構成する単語列の各単語に対応するタイムステップごとに1以上の単語を、対訳文に用いられる単語としての確からしさを示す確率と共に、順次出力する言語モデルを含む、ニューラル機械翻訳エンジンであり、生成部は、機械翻訳エンジンに原文を入力することにより、各タイムステップにおいて出力された1以上の単語、及び、各単語の対訳文を構成する単語としての確からしさを示す第1の確率に基づいて、複数の翻訳文を生成することとしてもよい。

[0085] 上記形態によれば、各タイムステップにおいて出力される複数の単語のそれぞれの第1の確率が参照されて翻訳文が生成されるので、取得された原文に対応する翻訳文としての尤度が高い複数の翻訳文を生成することができる。

[0086] また、別の形態に係る提示装置では、生成部は、ビームサーチの手法によ

り、対訳文を構成する単語列の各単語を、各単語の第1の確率に基づいて算出される単語列全体の対訳文としての確からしさを示す第2の確率に基づいて、文頭から順次抽出し、第2の確率の上位所定数の対訳文を、複数の翻訳文として出力することとしてもよい。

[0087] 上記形態によれば、文頭から順次抽出された単語により構成される単語列に関する第2の確率が算出され、第2の確率の上位所定数の対訳文が出力されるので、原文に対応する翻訳文としての尤度が高い単語列を翻訳文として出力できる。また、出力された複数の翻訳文に含まれる単語が提示されるので、多様な表現の翻訳文に対応した単語の候補の提示が可能となる。

[0088] また、別の形態に係る提示装置では、生成部は、T o p - K サンプリングの手法により、各タイムステップにおいて、第1の確率の上位所定数の単語を順次抽出し、抽出された単語からなる単語列により構成される複数の対訳文を、複数の翻訳文として出力することとしてもよい。

[0089] 上記形態によれば、各タイムステップにおいて第1の確率の上位所定数の単語が抽出され、抽出された単語に基づいて構成される対訳文が出力される。従って、原文に対応する翻訳文としての尤度が高い単語列を翻訳文として出力できる。また、各タイムステップにおいて抽出する単語の所定数を調整する（多くする）ことにより、誤りを含んだ単語の候補の提示が可能となる。

[0090] また、別の形態に係る提示装置では、機械翻訳エンジンは、エンコーダデコーダモデルを含み、提示装置は、提示部により提示された複数の単語のうちの少なくとも一つの単語を指定する指定入力を受け付ける指定入力受付部、を更に備え、生成部は、指定入力受付部により受け付けられた単語である指定単語をエンコーダデコーダモデルのデコーダに入力し、指定単語の入力に引き続いてデコーダの各タイムステップにおいて第1の確率に基づいて複数出力される単語を順次抽出し、順次抽出された単語からなる単語列に基づいて、複数の翻訳文を生成し、提示部は、指定入力受付部により指定単語が受け付けられるたびに生成部により生成された複数の翻訳文に含まれる単語

を提示することとしてもよい。

- [0091] 上記形態によれば、指定入力された単語からなる単語列を制約として、単語列に引き続く尤もらしい翻訳文が複数生成される。従って、単語列に引き続く単語の候補を、指定入力に応じて動的に変更して提示できる。
- [0092] また、別の形態に係る提示装置では、機械翻訳エンジンは、第1言語の文を構成する単語の第2言語による対訳の候補である単語対訳候補を、対訳としての確からしさを示す第3の確率と共に出力する翻訳モデルと、単語対訳候補の対訳文を構成する単語としての確からしさを示す第4の確率を出力する言語モデルとを含む統計的機械翻訳エンジンであり、生成部は、第3の確率及び第4の確率に基づいて、複数の翻訳文を生成することとしてもよい。
- [0093] 上記形態によれば、統計的機械翻訳により原文に対応する複数の翻訳文が生成される。従って、取得された原文に対応する翻訳文としての尤度が高い複数の翻訳文に含まれる単語を提示できる。
- [0094] また、別の形態に係る提示装置では、第1言語の原文により構成される問題文を提示する問題文提示部と、提示部により提示された単語に対する指定入力を受け付け、指定入力された単語を指定入力の順に応じて配列してなる作成文を生成し、作成文を提示する作文部と、をさらに備えることとしてもよい。
- [0095] 上記形態によれば、問題文の提示に応じて指定入力された単語の配列からなる作成文が生成及び提示される。これにより、原文に対応する翻訳の学習を実施できる。
- [0096] また、別の形態に係る提示装置では、提示部は、複数の翻訳文に含まれる単語をランダムに配列して提示することとしてもよい。
- [0097] 上記形態によれば、提示された単語の配列順に基づく、翻訳文における単語の配列順の認識を不可能にさせることができる。
- [0098] 以上、本実施形態について詳細に説明したが、当業者にとっては、本実施形態が本明細書中に説明した実施形態に限定されるものではないということは明らかである。本実施形態は、特許請求の範囲の記載により定まる本発明

の趣旨及び範囲を逸脱することなく修正及び変更態様として実施することができる。したがって、本明細書の記載は、例示説明を目的とするものであり、本実施形態に対して何ら制限的な意味を有するものではない。

[0099] 本明細書で説明した各態様／実施形態は、LTE (Long Term Evolution)、LTE-A (LTE-Advanced)、SUPER 3G、IMT-Advanced、4G、5G、FRA (Future Radio Access)、W-CDMA (登録商標)、GSM (登録商標)、CDMA2000、UMB (Ultra Mobile Broadband)、IEEE 802.11 (Wi-Fi)、IEEE 802.16 (WiMAX)、IEEE 802.20、UWB (Ultra-WideBand)、Bluetooth (登録商標)、その他の適切なシステムを利用するシステム及び／又はこれらに基づいて拡張された次世代システムに適用されてもよい。

[0100] 本明細書で説明した各態様／実施形態の処理手順、シーケンス、フローチャートなどは、矛盾の無い限り、順序を入れ替えてもよい。例えば、本明細書で説明した方法については、例示的な順序で様々なステップの要素を提示しており、提示した特定の順序に限定されない。

[0101] 入出力された情報等は特定の場所(例えば、メモリ)に保存されてもよいし、管理テーブルで管理してもよい。入出力される情報等は、上書き、更新、または追記され得る。出力された情報等は削除されてもよい。入力された情報等は他の装置へ送信されてもよい。

[0102] 判定は、1ビットで表される値(0か1か)によって行われてもよいし、真偽値(Boolean: trueまたはfalse)によって行われてもよいし、数値の比較(例えば、所定の値との比較)によって行われてもよい。

[0103] 本明細書で説明した各態様／実施形態は単独で用いてもよいし、組み合わせて用いてもよいし、実行に伴って切り替えて用いてもよい。また、所定の情報の通知(例えば、「Xであること」の通知)は、明示的に行うものに限られず、暗黙的(例えば、当該所定の情報の通知を行わない)ことによって行われてもよい。

[0104] 以上、本開示について詳細に説明したが、当業者にとっては、本開示が本開示中に説明した実施形態に限定されるものではないということは明らかである。本開示は、請求の範囲の記載により定まる本開示の趣旨及び範囲を逸脱することなく修正及び変更態様として実施することができる。したがって、本開示の記載は、例示説明を目的とするものであり、本開示に対して何ら制限的な意味を有するものではない。

[0105] ソフトウェアは、ソフトウェア、ファームウェア、ミドルウェア、マイクロコード、ハードウェア記述言語と呼ばれるか、他の名称で呼ばれるかを問わず、命令、命令セット、コード、コードセグメント、プログラムコード、プログラム、サブプログラム、ソフトウェアモジュール、アプリケーション、ソフトウェアアプリケーション、ソフトウェアパッケージ、ルーチン、サブルーチン、オブジェクト、実行可能ファイル、実行スレッド、手順、機能などを意味するよう広く解釈されるべきである。

[0106] また、ソフトウェア、命令などは、伝送媒体を介して送受信されてもよい。例えば、ソフトウェアが、同軸ケーブル、光ファイバケーブル、ツイストペア及びデジタル加入者回線（DSL）などの有線技術及び／又は赤外線、無線及びマイクロ波などの無線技術を使用してウェブサイト、サーバ、又は他のリモートソースから送信される場合、これらの有線技術及び／又は無線技術は、伝送媒体の定義内に含まれる。

[0107] 本開示において説明した情報、信号などは、様々な異なる技術のいずれかを使用して表されてもよい。例えば、上記の説明全体に渡って言及され得るデータ、命令、コマンド、情報、信号、ビット、シンボル、チップなどは、電圧、電流、電磁波、磁界若しくは磁性粒子、光場若しくは光子、又はこれらの任意の組み合わせによって表されてもよい。

[0108] なお、本開示において説明した用語及び／又は本明細書の理解に必要な用語については、同一の又は類似する意味を有する用語と置き換えてもよい。

[0109] 本明細書で使用する「システム」および「ネットワーク」という用語は、互換的に使用される。

[0110] また、本明細書で説明した情報、パラメータなどは、絶対値で表されてもよいし、所定の値からの相対値で表されてもよいし、対応する別の情報で表されてもよい。

[0111] 本開示で使用する「判断(determining)」、「決定(determining)」という用語は、多種多様な動作を包含する場合がある。「判断」、「決定」は、例えば、判定(judging)、計算(calculating)、算出(computing)、処理(processing)、導出(deriving)、調査(investigating)、探索(looking up、search、inquiry) (例えば、テーブル、データベース又は別のデータ構造での探索)、確認(ascertaining)した事を「判断」「決定」したとみなす事などを含み得る。また、「判断」、「決定」は、受信(receiving) (例えば、情報を受信すること)、送信(transmitting) (例えば、情報を送信すること)、入力(input)、出力(output)、アクセス(accessing) (例えば、メモリ中のデータにアクセスすること)した事を「判断」「決定」したとみなす事などを含み得る。また、「判断」、「決定」は、解決(resolving)、選択(selecting)、選定(choosing)、確立(establishing)、比較(comparing)などした事を「判断」「決定」したとみなす事を含み得る。つまり、「判断」「決定」は、何らかの動作を「判断」「決定」したとみなす事を含み得る。また、「判断(決定)」は、「想定する(assuming)」、「期待する(expecting)」、「みなす(considering)」などで読み替えられてもよい。

[0112] 本開示で使用する「に基づいて」という記載は、別段に明記されていない限り、「のみに基づいて」を意味しない。言い換えれば、「に基づいて」という記載は、「のみに基づいて」と「に少なくとも基づいて」の両方を意味する。

[0113] 本明細書で「第1の」、「第2の」などの呼称を使用した場合においては、その要素へのいかなる参照も、それらの要素の量または順序を全般的に限定するものではない。これらの呼称は、2つ以上の要素間を区別する便利な方法として本明細書で使用され得る。したがって、第1および第2の要素への参照は、2つの要素のみがそこで採用され得ること、または何らかの形で

第1の要素が第2の要素に先行しなければならないことを意味しない。

[0114] 「含む(include)」、「含んでいる(including)」、およびそれらの変形が、本明細書あるいは特許請求の範囲で使用されている限り、これら用語は、用語「備える(comprising)」と同様に、包括的であることが意図される。さらに、本明細書あるいは特許請求の範囲において使用されている用語「または(or)」は、排他的論理和ではないことが意図される。

[0115] 本明細書において、文脈または技術的に明らかに1つのみしか存在しない装置である場合以外は、複数の装置をも含むものとする。

[0116] 本開示の全体において、文脈から明らかに単数を示したものではなくれば、複数のものを含むものとする。

符号の説明

[0117] 10…提示装置、11…取得部、12…生成部、13…提示部、14…指定入力受付部、15…問題文提示部、16…作文部、20…原文記憶部、30…端末、M1…記録媒体、m10…メインモジュール、m11…取得モジュール、m12…生成モジュール、m13…提示モジュール、m14…指定入力受付モジュール、m15…問題文提示モジュール、m16…作文モジュール、P1…提示プログラム。

請求の範囲

- [請求項1] 第1言語からなる原文の第2言語からなる翻訳文を構成する単語の候補を提示する提示装置であって、
- 入力された第1言語の文に基づいて第2言語からなる対訳文を出力する機械翻訳エンジンに前記原文を入力することにより前記機械翻訳エンジンから出力される対訳文に基づいて、第2言語からなる複数の翻訳文を生成する生成部と、
- 前記複数の翻訳文に含まれる単語を、所定の配列手法により配列して提示する提示部と、
- を備える提示装置。
- [請求項2] 前記機械翻訳エンジンは、ニューラルネットワークを含み、第1言語の文の入力に基づいて、前記対訳文を構成する単語列の各単語に対応するタイムステップごとに1以上の単語を、前記対訳文に用いられる単語としての確からしさを示す確率と共に、順次出力する言語モデルを含む、ニューラル機械翻訳エンジンであり、
- 前記生成部は、前記機械翻訳エンジンに前記原文を入力することにより、各タイムステップにおいて出力された1以上の単語、及び、各単語の前記対訳文を構成する単語としての確からしさを示す第1の確率に基づいて、前記複数の翻訳文を生成する、
- 請求項1に記載の提示装置。
- [請求項3] 前記生成部は、
- ビームサーチの手法により、前記対訳文を構成する単語列の各単語を、各単語の前記第1の確率に基づいて算出される単語列全体の前記対訳文としての確からしさを示す第2の確率に基づいて、文頭から順次抽出し、
- 前記第2の確率の上位所定数の対訳文を、前記複数の翻訳文として出力する、
- 請求項2に記載の提示装置。

- [請求項4] 前記生成部は、
Top-Kサンプリングの手法により、各タイムステップにおいて、前記第1の確率の上位所定数の単語を順次抽出し、
抽出された単語からなる単語列により構成される複数の対訳文を、前記複数の翻訳文として出力する、
請求項2に記載の提示装置。
- [請求項5] 前記機械翻訳エンジンは、エンコーダデコーダモデルを含み、
前記提示装置は、
前記提示部により提示された複数の単語のうちの少なくとも一つの単語を指定する指定入力を受け付ける指定入力受付部、を更に備え、
前記生成部は、前記指定入力受付部により受け付けられた単語である指定単語を前記エンコーダデコーダモデルのデコーダに入力し、前記指定単語の入力に引き続いて前記デコーダの各タイムステップにおいて前記第1の確率に基づいて複数出力される単語を順次抽出し、順次抽出された単語からなる単語列に基づいて、前記複数の翻訳文を生成し、
前記提示部は、前記指定入力受付部により前記指定単語が受け付けられるたびに前記生成部により生成された前記複数の翻訳文に含まれる単語を提示する、
請求項2～4のいずれか一項に記載の提示装置。
- [請求項6] 前記機械翻訳エンジンは、第1言語の文を構成する単語の第2言語による対訳の候補である単語対訳候補を、対訳としての確からしさを示す第3の確率と共に出力する翻訳モデルと、前記単語対訳候補の前記対訳文を構成する単語としての確からしさを示す第4の確率を出力する言語モデルとを含む統計的機械翻訳エンジンであり、
前記生成部は、前記第3の確率及び前記第4の確率に基づいて、前記複数の翻訳文を生成する、
請求項1に記載の提示装置。

[請求項7] 第1言語の前記原文により構成される問題文を提示する問題文提示部と、

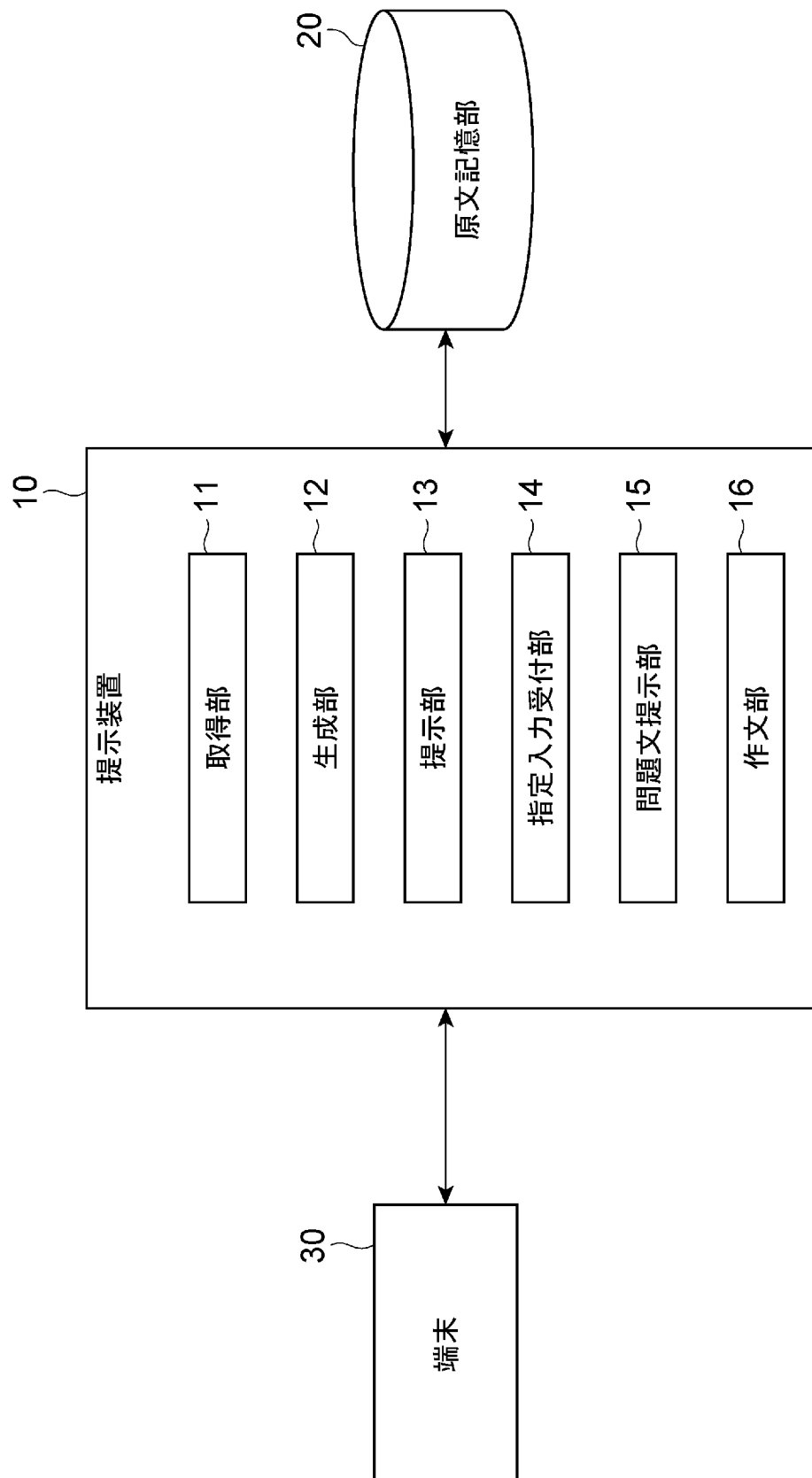
前記提示部により提示された単語に対する指定入力を受け付け、指定入力された単語を指定入力の順に応じて配列してなる作成文を生成し、前記作成文を提示する作文部と、をさらに備える、

請求項1～6のいずれか一項に記載の提示装置。

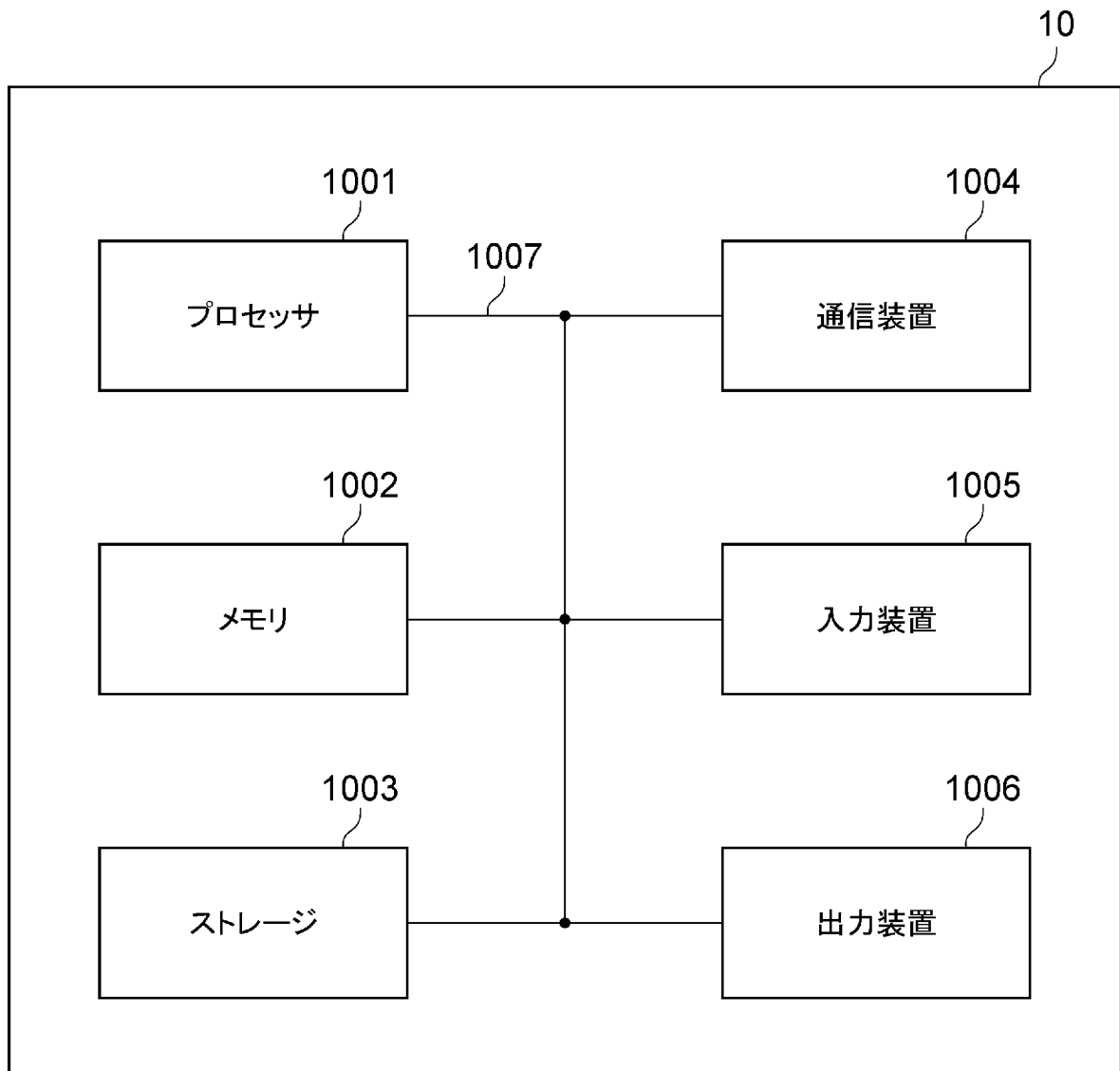
[請求項8] 前記提示部は、前記複数の翻訳文に含まれる単語をランダムに配列して提示する、

請求項1～7のいずれか一項に記載の提示装置。

[図1]



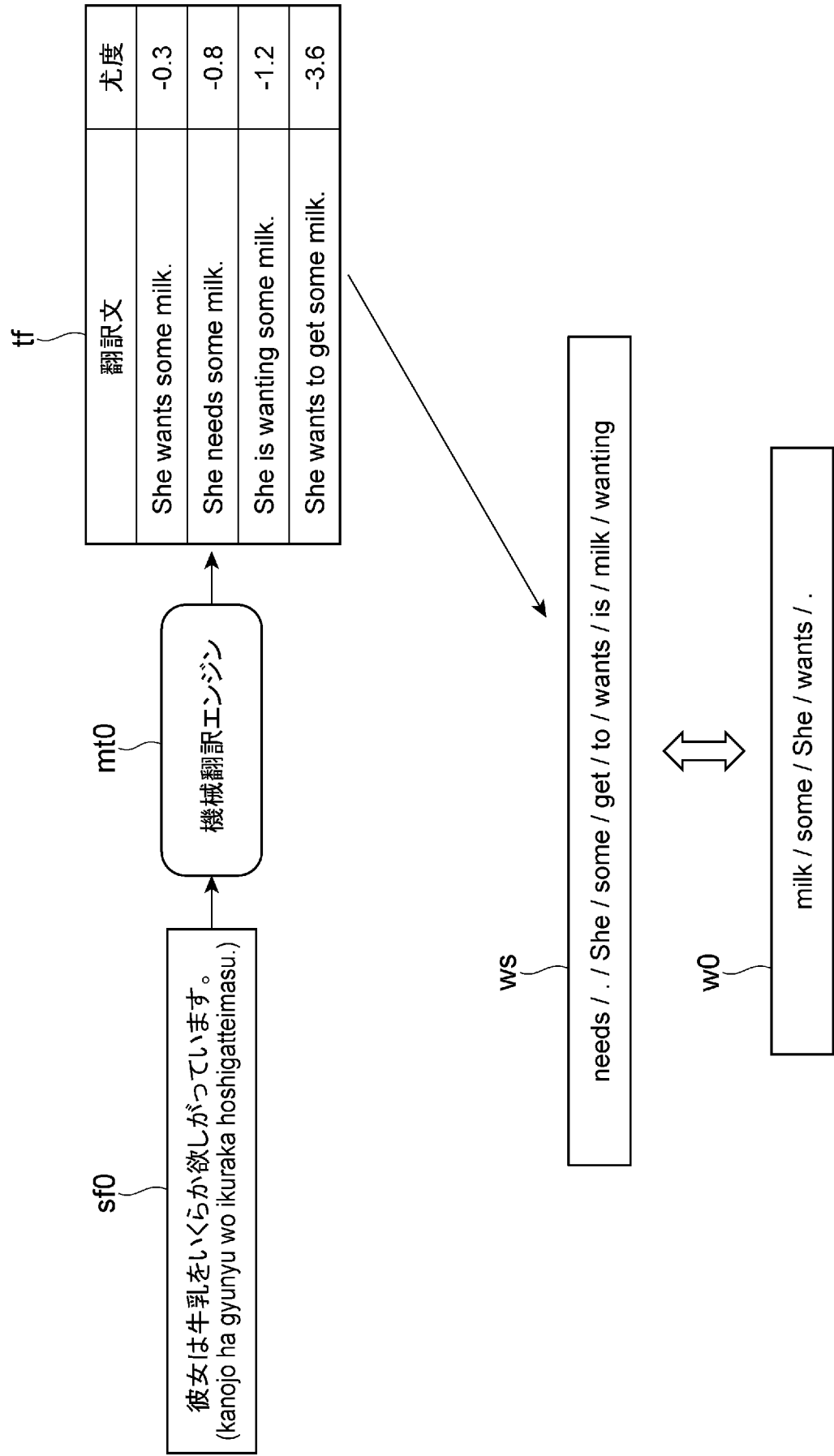
[図2]



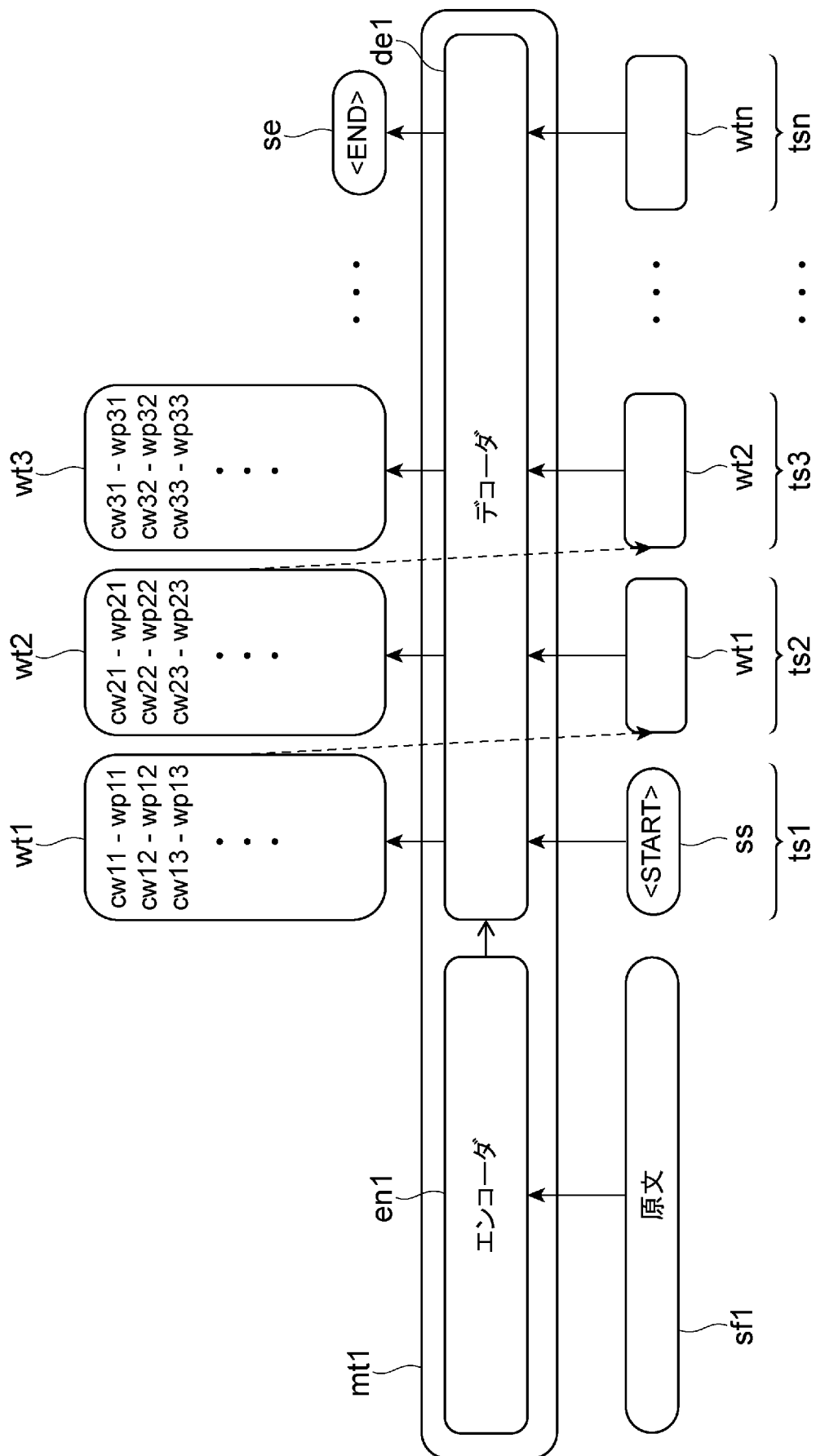
[図3]

ID	原文
1	彼女は牛乳をいくらかほしがっています。 (kanojo ha gyunyu wo ikuraka hoshigatteimasu.)
2	私は公園の中を歩いています。 (watashi ha koen no naka wo hashitteimasu.)
3	彼はサッカーの試合でけがをした。 (kare ha sakka no shiai de kega wo shita.)
4	私はあなたに幸せなままに続けてほしい。 (watashi ha anata ni shiawasenamamade itehoshii.)
⋮	⋮

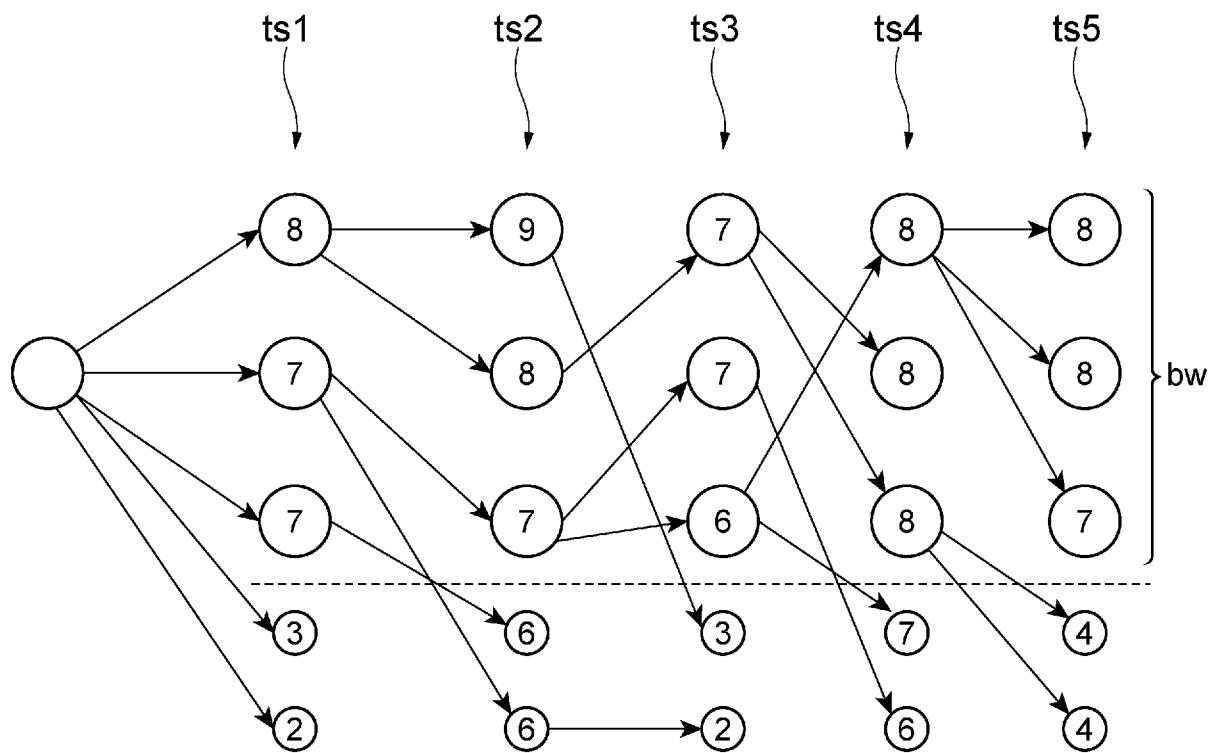
[図4]



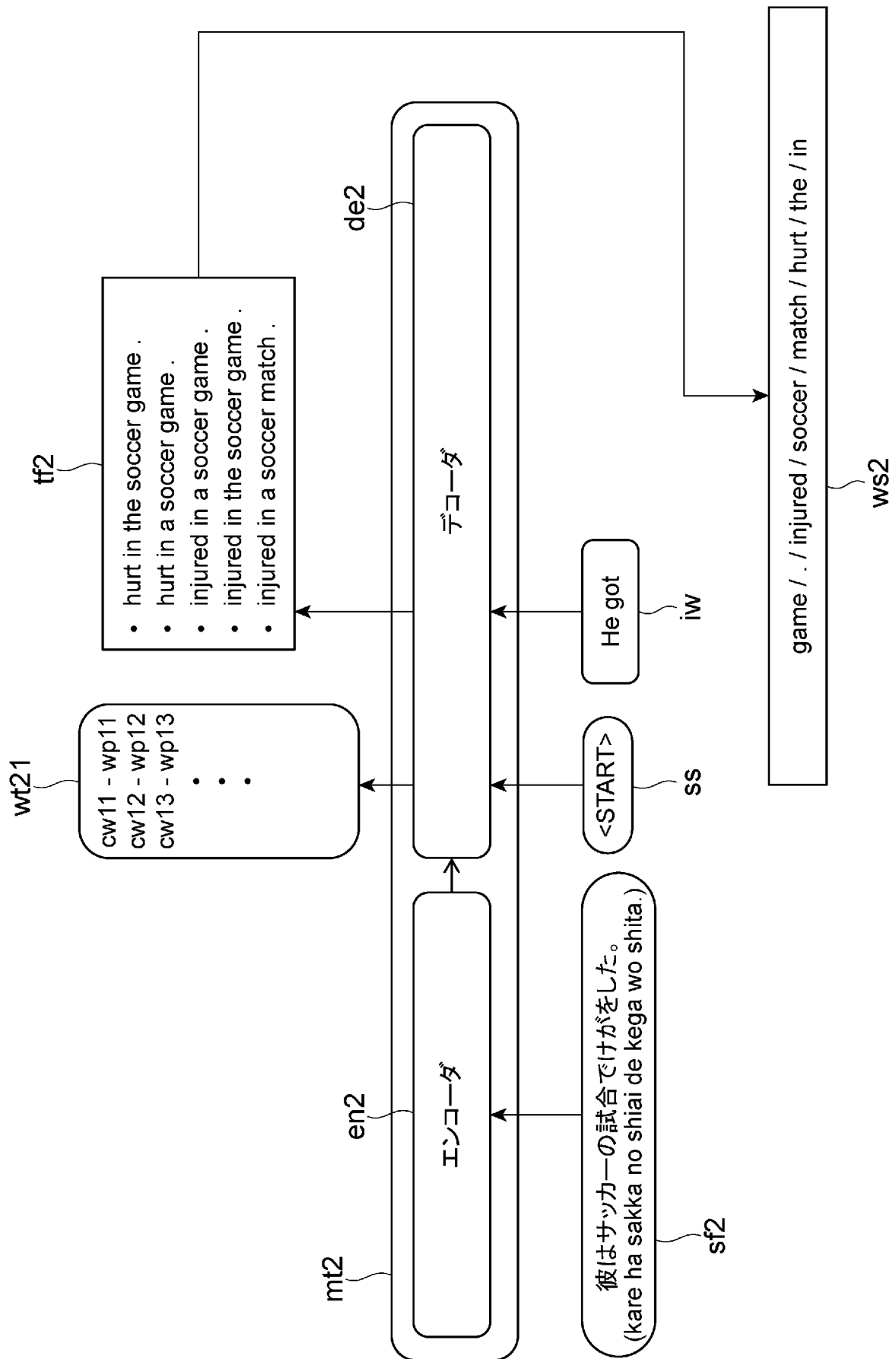
[図5]



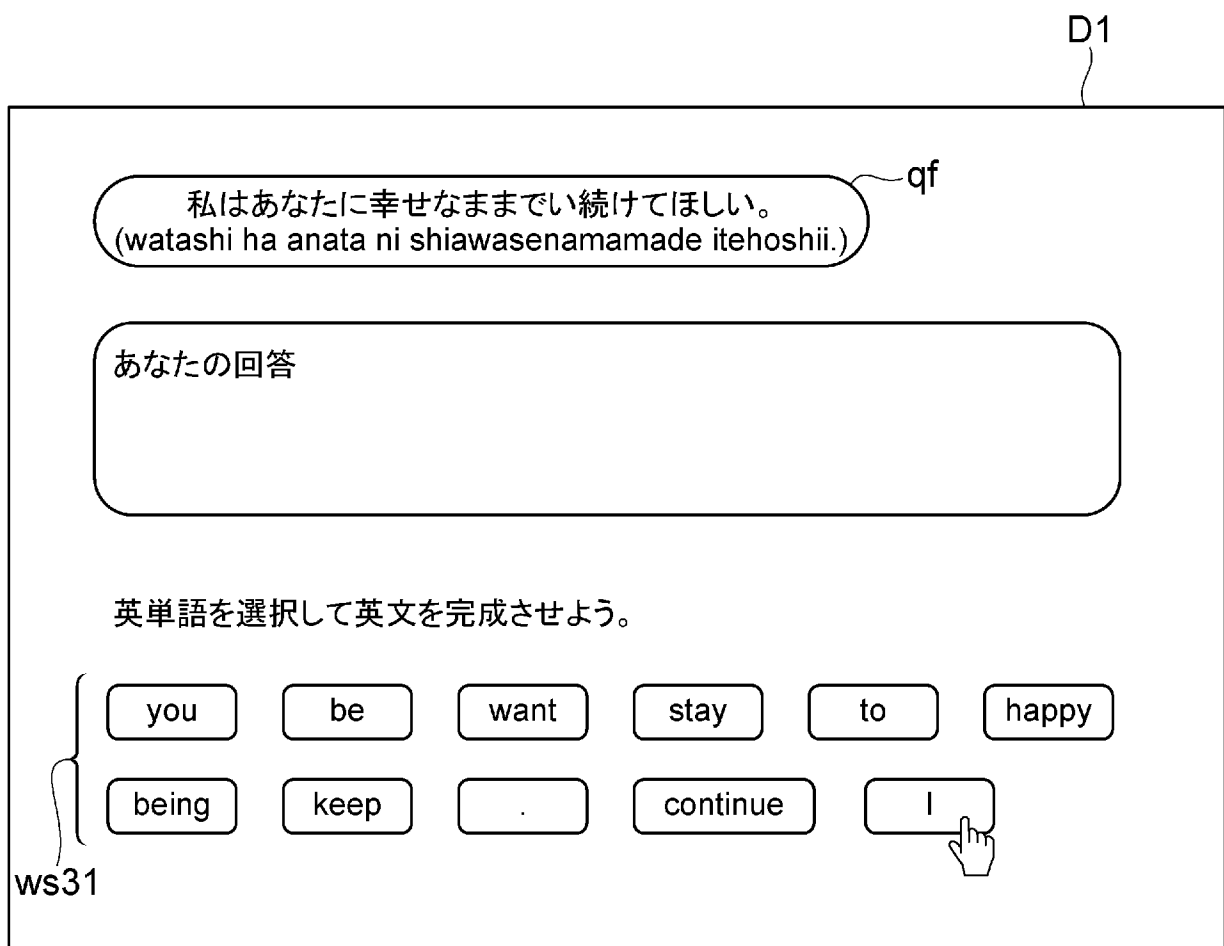
[図6]



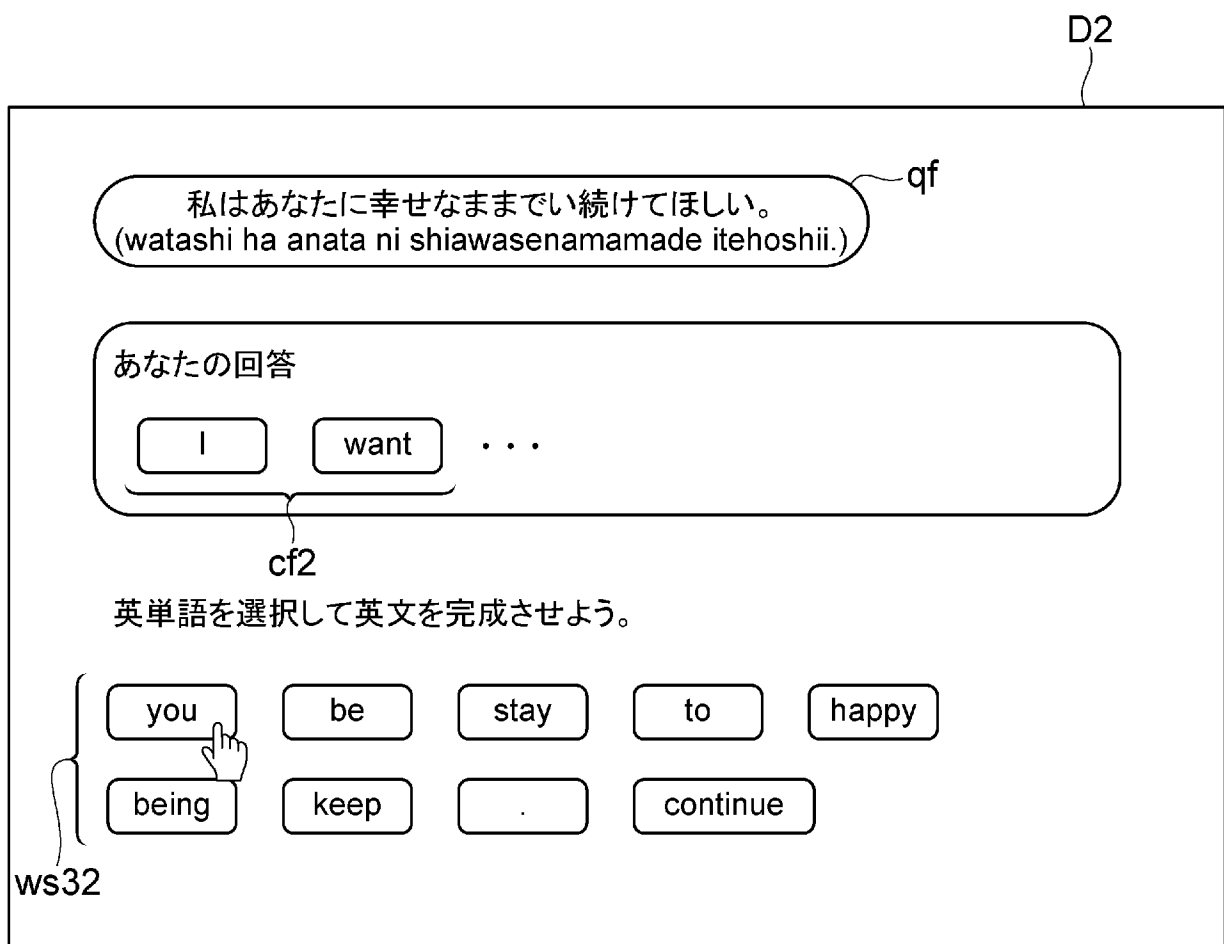
[図7]



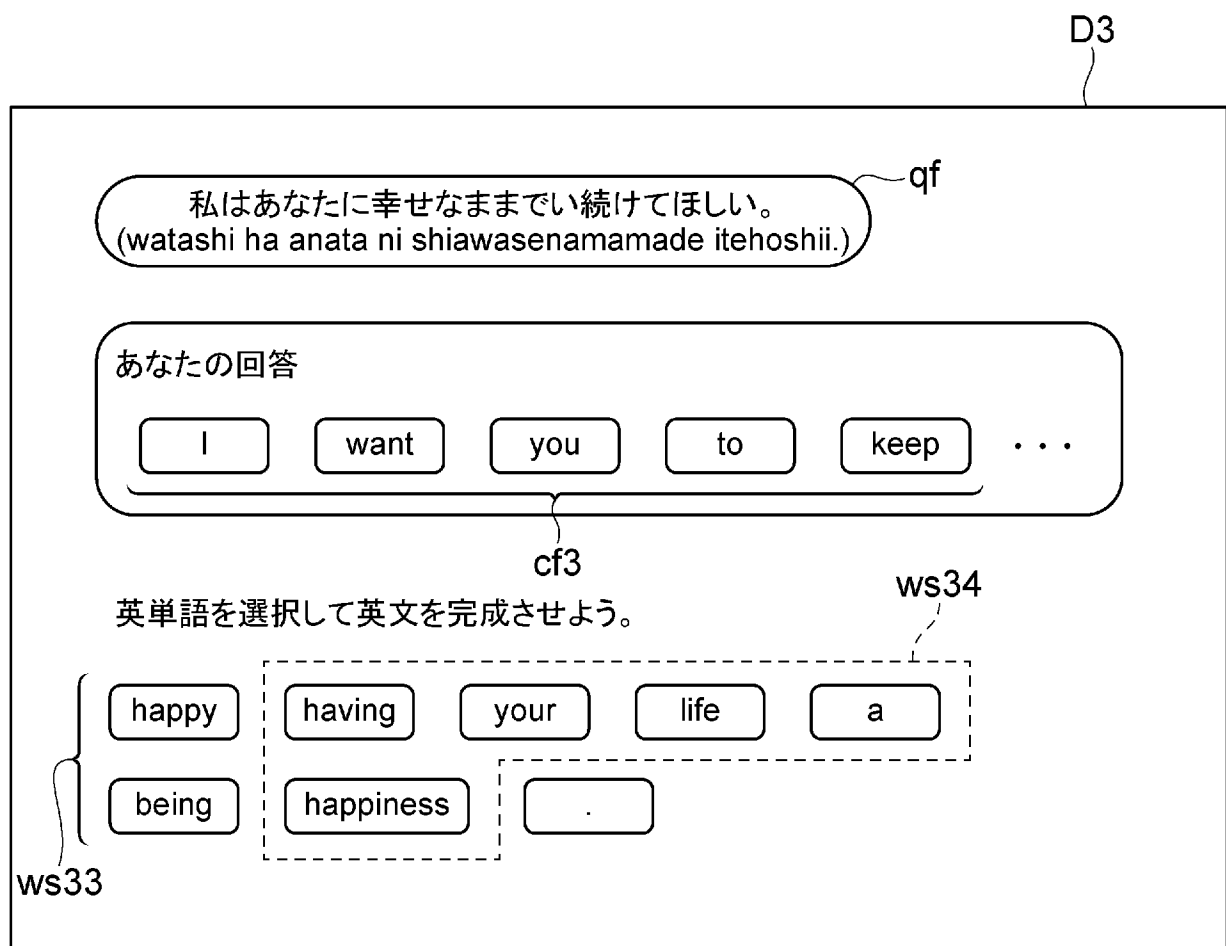
[図8]



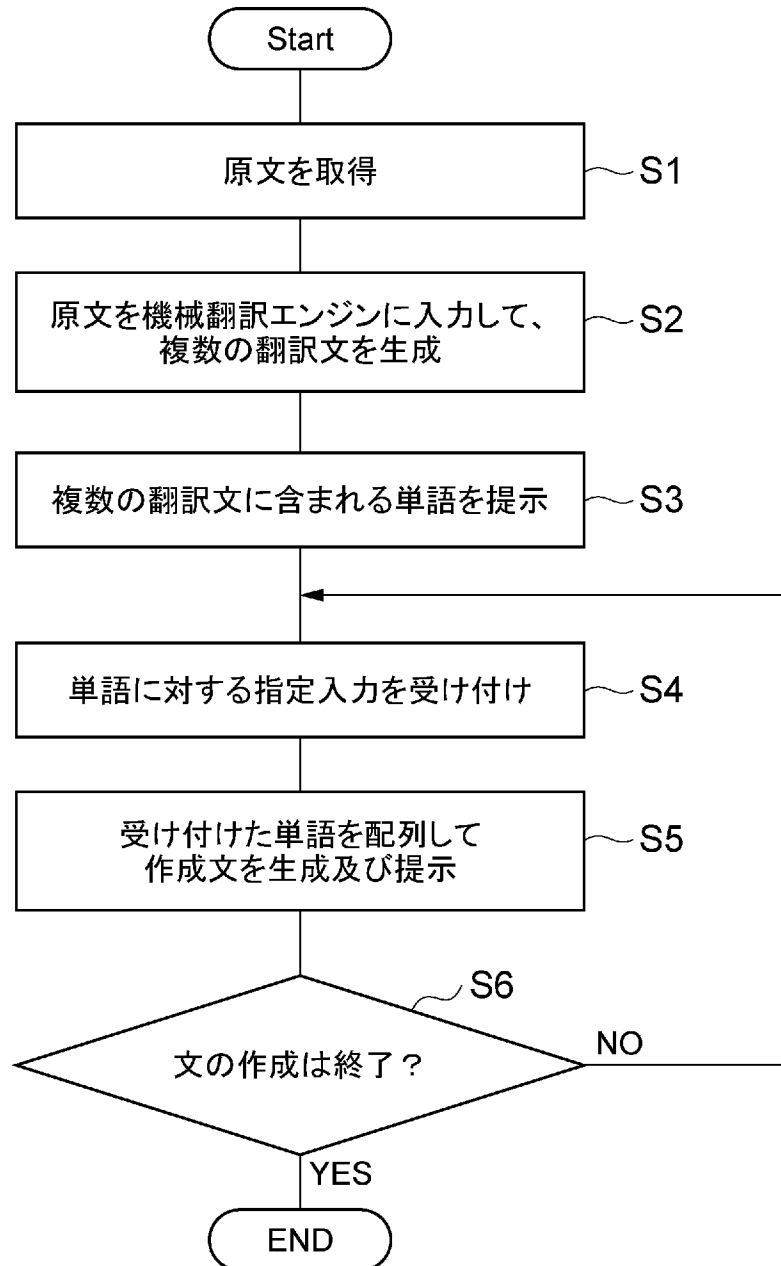
[図9]



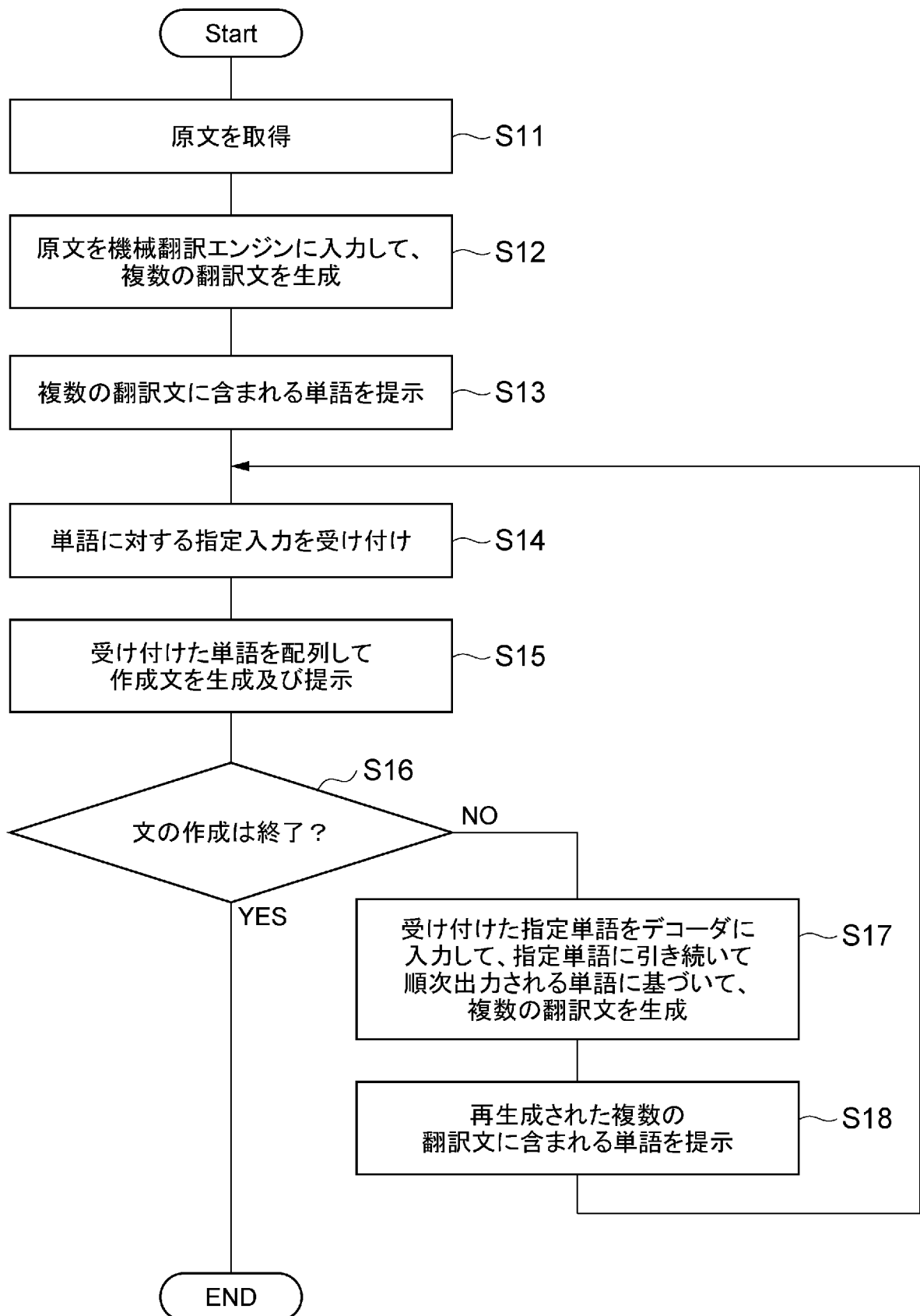
[図10]



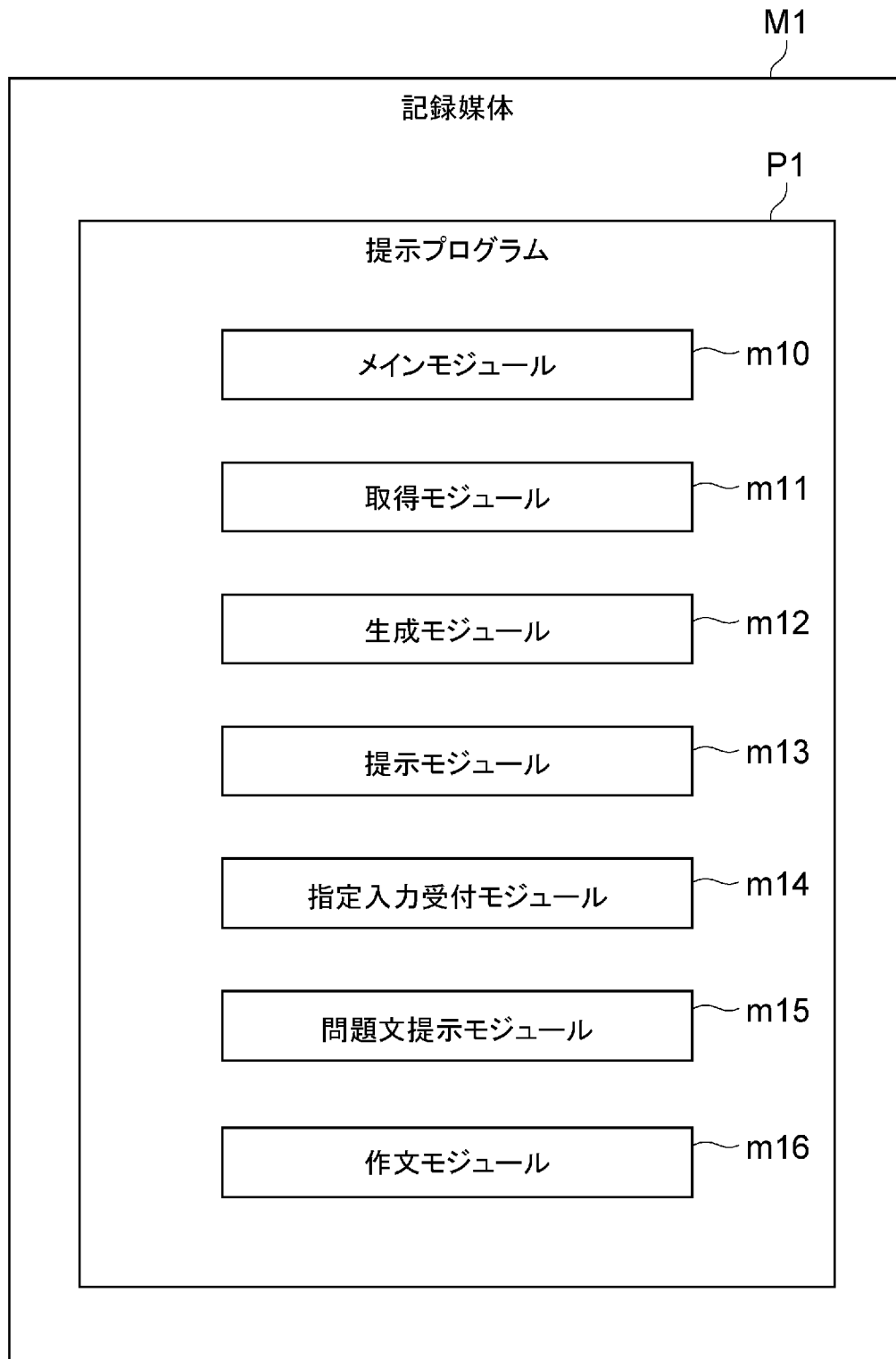
[図11]



[図12]



[図13]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2021/043451

A. CLASSIFICATION OF SUBJECT MATTER**G06F 40/44**(2020.01)i; **G06F 40/47**(2020.01)i; **G06F 40/58**(2020.01)i

FI: G06F40/47; G06F40/44; G06F40/58

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)

G06F40/44; G06F40/47; G06F40/58; G09B19/06; G06B7/04

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 1971-2022
 Registered utility model specifications of Japan 1996-2022
 Published registered utility model applications of Japan 1994-2022

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.
Y	JP 2008-197604 A (WATANABE, Shoichi) 28 August 2008 (2008-08-28) paragraphs [0005], [0011], fig. 3, 4	1-8
Y	CN 1770137 A (INVENTEC CORP.) 10 May 2006 (2006-05-10) p. 8, lines 15-21, p. 9, lines 11-22, fig. 3b-3d	1-8
Y	JP 2012-118883 A (NEC CORP.) 21 June 2012 (2012-06-21) paragraphs [0008], [0052], [0062]-[0067], [0070]-[0072], [0074], [0075]	1-8
Y	JP 2018-5218 A (SAMSUNG ELECTRONICS CO., LTD.) 11 January 2018 (2018-01-11) paragraphs [0088], [0089], [0094], [0154]	2-5, 7, 8
Y	WO 2019/225154 A1 (NTT DOCOMO, INC.) 28 November 2019 (2019-11-28) paragraphs [0017], [0025]-[0027], fig. 2	2-5, 7, 8
Y	CN 111507113 A (BEIJING SINOVOICE TECHNOLOGY CO., LTD.) 07 August 2020 (2020-08-07) paragraphs [0057]-[0064]	2-5, 7, 8

☐ 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

“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

17 February 2022

Date of mailing of the international search report

01 March 2022

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/JP2021/043451

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP	2008-197604	A	28 August 2008	(Family: none)	
CN	1770137	A	10 May 2006	(Family: none)	
JP	2012-118883	A	21 June 2012	(Family: none)	
JP	2018-5218	A	11 January 2018	US 2018/0011843 A1 paragraphs [0091], [0092], [0095], [0132] EP 3267328 A1 KR 10-2018-0005850 A CN 107590135 A	
WO	2019/225154	A1	28 November 2019	US 2021/0019478 A1 paragraphs [0021], [0029]- [0031], fig. 2	
CN	111507113	A	07 August 2020	(Family: none)	

A. 発明の属する分野の分類（国際特許分類（IPC）） G06F 40/44(2020.01)i; G06F 40/47(2020.01)i; G06F 40/58(2020.01)i FI: G06F40/47; G06F40/44; G06F40/58														
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）） G06F40/44; G06F40/47; G06F40/58, G09B19/06; G06B7/04														
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 <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922 - 1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971 - 2022年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996 - 2022年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994 - 2022年</td> </tr> </table>			日本国実用新案公報	1922 - 1996年	日本国公開実用新案公報	1971 - 2022年	日本国実用新案登録公報	1996 - 2022年	日本国登録実用新案公報	1994 - 2022年				
日本国実用新案公報	1922 - 1996年													
日本国公開実用新案公報	1971 - 2022年													
日本国実用新案登録公報	1996 - 2022年													
日本国登録実用新案公報	1994 - 2022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												
Y	JP 2008-197604 A (渡辺 昇一) 28.08.2008 (2008 - 08 - 28) 段落0005, 0011、図3, 4	1-8												
Y	CN 1770137 A (INVENTEC CORPORATION) 10.05.2006 (2006 - 05 - 10) 第8頁第15-21行、第9頁第11-22行、図3b-3d	1-8												
Y	JP 2012-118883 A (日本電気株式会社) 21.06.2012 (2012 - 06 - 21) 段落0008, 0052, 0062-0067, 0070-0072, 0074-0075	1-8												
Y	JP 2018-5218 A (三星電子株式会社) 11.01.2018 (2018 - 01 - 11) 段落0088, 0089, 0094, 0154	2-5, 7, 8												
Y	WO 2019/225154 A1 (株式会社NTTドコモ) 28.11.2019 (2019 - 11 - 28) 段落0017, 0025-0027、図2	2-5, 7, 8												
Y	CN 111507113 A (BEIJING SINOVOICE TECHNOLOGY CO., LTD.) 07.08.2020 (2020 - 08 - 07) 段落0057-0064	2-5, 7, 8												
☐ 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。														
<table border="0"> <tr> <td>* 引用文献のカテゴリー</td> <td>“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</td> </tr> <tr> <td>“A” 特に関連のある文献ではなく、一般的な技術水準を示すもの</td> <td>“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</td> </tr> <tr> <td>“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</td> <td>“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</td> </tr> <tr> <td>“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）</td> <td>“&” 同一パテントファミリー文献</td> </tr> <tr> <td>“O” 口頭による開示、使用、展示等に言及する文献</td> <td></td> </tr> <tr> <td>“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</td> <td></td> </tr> </table>			* 引用文献のカテゴリー	“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	“A” 特に関連のある文献ではなく、一般的な技術水準を示すもの	“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	“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	“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	“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）	“&” 同一パテントファミリー文献	“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	
* 引用文献のカテゴリー	“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													
“A” 特に関連のある文献ではなく、一般的な技術水準を示すもの	“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													
“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	“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													
“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）	“&” 同一パテントファミリー文献													
“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														
国際調査を完了した日 17.02.2022	国際調査報告の発送日 01.03.2022													
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） 長 由紀子 5N 4233 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/JP2021/043451

引用文献			公表日	パテントファミリー文献	公表日
JP	2008-197604	A	28.08.2008	(ファミリーなし)	
CN	1770137	A	10.05.2006	(ファミリーなし)	
JP	2012-118883	A	21.06.2012	(ファミリーなし)	
JP	2018-5218	A	11.01.2018	US 2018/0011843 A1	
				段落0091, 0092, 0095, 0132	
				EP 3267328 A1	
				KR 10-2018-0005850 A	
				CN 107590135 A	
WO	2019/225154	A1	28.11.2019	US 2021/0019478 A1	
				段落0021, 0029-0031、図2	
CN	111507113	A	07.08.2020	(ファミリーなし)	