

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第6675463号
(P6675463)

(45) 発行日 令和2年4月1日 (2020. 4. 1)

(24) 登録日 令和2年3月12日 (2020. 3. 12)

(51) Int. Cl.

F I

G 1 O L 15/22 (2006. 01)

G 1 O L 15/197 (2013. 01)

G 1 O L 15/18 (2013. 01)

G 1 O L 15/22 4 7 O Z

G 1 O L 15/197

G 1 O L 15/18 3 O O G

請求項の数 20 外国語出願 (全 35 頁)

(21) 出願番号	特願2018-189730 (P2018-189730)	(73) 特許権者	507134068
(22) 出願日	平成30年10月5日 (2018. 10. 5)		サウンドハウンド、インコーポレイテッド
(65) 公開番号	特開2019-70799 (P2019-70799A)		アメリカ合衆国、9 5 0 5 4 カリフォル
(43) 公開日	令和1年5月9日 (2019. 5. 9)		ニア州、サンタ・クララ、ベッツィー・ロ
審査請求日	令和1年9月11日 (2019. 9. 11)		ス・ドライブ、5 4 0 0
(31) 優先権主張番号	15/726, 394	(74) 代理人	110001195
(32) 優先日	平成29年10月6日 (2017. 10. 6)		特許業務法人深見特許事務所
(33) 優先権主張国・地域又は機関	米国 (US)	(72) 発明者	ルーク・レフェビュア
			アメリカ合衆国、9 5 0 5 4 カリフォル
			ニア州、サンタ・クララ、ベッツィー・ロ
			ス・ドライブ、5 4 0 0
早期審査対象出願		(72) 発明者	プラナフ・シン
			アメリカ合衆国、9 5 0 5 4 カリフォル
			ニア州、サンタ・クララ、ベッツィー・ロ
			ス・ドライブ、5 4 0 0
			最終頁に続く

(54) 【発明の名称】 自然言語の双方向確率的な書換えおよび選択

(57) 【特許請求の範囲】

【請求項 1】

クエリ結果をユーザに提供する際に、ユーザクエリの入力トークンシーケンスを書き換えるためのコンピュータ実装方法であって、

コンピュータネットワークを介して、遠隔クライアント装置上で動作するアプリケーションから、システムユーザからのユーザクエリを受信するステップを含み、前記ユーザクエリは、前記入力トークンシーケンスを含み、

正方向統計言語モデルに従って、前記入力トークンシーケンス内の複数のトークンの正方向確率を決定するステップと、

反方向統計言語モデルに従って、前記入力トークンシーケンス内の複数のトークンの反方向確率を決定するステップと、

低い反方向確率を有する第 1 のトークンの後方且つ低い正方向確率を有する隣接の第 2 のトークンの前方の位置に新たなトークンを挿入することによって、新たな書換えトークンシーケンスを作成するステップと、

前記新たな書換えトークンシーケンスを処理することによって、結果を生成するステップと、

前記処理による前記結果を示す応答を前記システムユーザに提供するステップとを含む、方法。

【請求項 2】

入力トークンシーケンスを書き換えるためのコンピュータ実装方法であって、

正方向統計言語モデルに従って、前記入力トークンシーケンス内の複数のトークンの正方向確率を決定するステップと、

反方向統計言語モデルに従って、前記入力トークンシーケンス内の複数のトークンの反方向確率を決定するステップと、

前記複数のトークンの各トークンの確率スコアを前記正方向確率のうち最も低いものと前記反方向確率のうち最も低いものとに基づいて計算するステップと、

新たなトークンを用いて、最も低いスコアを有するトークンを置換することによって、新たな書換えトークンシーケンスを作成するステップとを含む、方法。

【請求項 3】

入力トークンシーケンスを書き換えるためのコンピュータ実装方法であって、

正方向統計言語モデル (S L M) に従って、前記入力トークンシーケンス内の複数のトークンの正方向確率を決定するステップと、

反方向統計言語モデルに従って、前記入力トークンシーケンス内の複数のトークンの反方向確率を決定するステップと、

新たなトークンを用いて、第 1 の閾値未満の正方向確率且つ第 2 の閾値未満の反方向確率を有する疑わしいトークンを置換することによって、新たな書換えトークンシーケンスを生成するステップとを含む、方法。

【請求項 4】

確率を決定する前に、タグを用いて、前記入力トークンシーケンス内の少なくとも 1 つのトークンを置換するステップをさらに含む、請求項 3 に記載の方法。

【請求項 5】

前記正方向 S L M に従って得られた最も高い確率を有するトークンのリストおよび前記反方向 S L M に従って得られた最も高い確率を有するトークンのリストの両方に存在するトークンを前記新たなトークンとして選択するステップをさらに含む、請求項 3 または 4 に記載の方法。

【請求項 6】

構文規則に従って、前記入力トークンシーケンスの統語解析を実行するステップと、

前記構文規則に従って、前記選択を隣接トークンの文脈において構文上合法的なトークンのみに制限するステップとをさらに含む、請求項 5 に記載の方法。

【請求項 7】

前記新たな書換えトークンシーケンスに対して、少なくとも前記正方向 S L M における前記新たなトークンの確率および前記反方向 S L M における前記新たなトークンの前記確率に依存する書換えスコアを計算するステップをさらに含む、請求項 5 または 6 に記載の方法。

【請求項 8】

多種多様なトピックに関連する表現から構築された多様コーパス S L M における前記新たなトークンの前記確率に基づいて、前記書換えスコアをスケールリングする、請求項 7 に記載の方法。

【請求項 9】

代替の新たなトークンを用いて、前記疑わしいトークンを置換することによって、代替の書換えトークンシーケンスを作成するステップと、

前記代替の書換えトークンシーケンスの代替スコアを、前記正方向 S L M における前記代替の新たなトークンの確率および前記反方向 S L M における前記代替の新たなトークンの確率の両方の組み合わせとして計算するステップと、

前記新たな書換えトークンシーケンスと前記代替の書換えトークンシーケンスとのうち、より高いスコアを有するいずれか一方を選択するステップとをさらに含む、請求項 5 ~ 8 のいずれか 1 項に記載の方法。

【請求項 10】

最近の連続音声から仮定されたトークンのトークンバッファを維持するステップをさらに含む、

10

20

30

40

50

前記入力トークンシーケンスは、前記トークンバッファ内のトークンからなるシーケンスである、請求項 5 ~ 9 のいずれか 1 項に記載の方法。

【請求項 1 1】

最近のトークンシーケンスに存在するトークンの履歴キャッシュを記憶するステップと、

、
正方向確率が最も高いトークンのリストおよび反方向確率が最も高いトークンのリストの各々から、前記疑わしいトークンを置換するための前記新たなトークンを選択するステップと、

前記履歴キャッシュに存在する少なくとも 1 つのトークンの前記確率スコアを増加させるステップとをさらに含む、請求項 3 ~ 1 0 のいずれか 1 項に記載の方法。

10

【請求項 1 2】

自然言語パーサを用いて、文法に従って前記新たな書換えトークンシーケンスを構文解析することによって、解析スコアを生成するステップをさらに含む、請求項 3 ~ 1 1 のいずれか 1 項に記載の方法。

【請求項 1 3】

前記疑わしいトークンとは異なる代替の疑わしいトークンを置換することによって、代替の書換えトークンシーケンスを作成するステップと、

自然言語パーサを用いて、前記文法に従って前記代替の書換えトークンシーケンスを構文解析することによって、代替の解析スコアを生成するステップと、

前記新たな書換えトークンシーケンスと前記代替の書換えトークンシーケンスとのうち、より高い解析スコアを有するいずれか一方を選択するステップとをさらに含む、請求項 1 2 に記載の方法。

20

【請求項 1 4】

代替の新たなトークンを用いて、前記疑わしいトークンを置換することによって、代替の書換えトークンシーケンスを作成するステップと、

自然言語パーサを用いて、前記文法に従って前記代替の書換えトークンシーケンスを構文解析することによって、代替の解析スコアを生成するステップと、

前記新たな書換えトークンシーケンスと前記代替の書換えトークンシーケンスとのうち、より高い解析スコアを有するいずれか一方を選択するステップとをさらに含む、請求項 1 2 または 1 3 に記載の方法。

30

【請求項 1 5】

ツリー型アルゴリズムを用いて、書換えを反復的に行い、各書換のスコアを計算することによって、書換のセットを生成し、前記書換のセットから最高のスコアを有する書換えを選択するステップをさらに含む、請求項 1 2 ~ 1 4 のいずれか 1 項に記載の方法。

【請求項 1 6】

前記入力トークンシーケンスをキャッシュに格納するステップと、

前記入力トークンシーケンスに関連して、前記新たな書換えトークンシーケンスを前記キャッシュに格納するステップと、

前記キャッシュを検索することによって、前記入力トークンシーケンスを探し出すステップとをさらに含む、請求項 3 ~ 1 5 のいずれか 1 項に記載の方法。

40

【請求項 1 7】

前記キャッシュを分析することによって、前記新たな書換えトークンシーケンスに対して、前記新たな書換えトークンシーケンスに書換えられた最も頻繁な入力トークンシーケンスを特定するステップをさらに含む、請求項 1 6 に記載の方法。

【請求項 1 8】

前記新たな書換えトークンシーケンスに書換えられた前記最も頻繁な入力トークンシーケンスを含むように、文法規則を作成するステップをさらに含む、請求項 1 7 に記載の方法。

【請求項 1 9】

50

前記新たな書換えトークンシーケンスに書換えられた前記最も頻繁な入力トークンシーケンスを含むように、文法規則を改変するステップをさらに含む、請求項 17 または 18 に記載の方法。

【請求項 20】

コンピュータプロセッサによって実行されるコンピュータプログラムであって、前記コンピュータプロセッサによって実行されると、前記コンピュータプロセッサに、請求項 1 ~ 19 のいずれか 1 項に記載の方法を実行させる、コンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

10

発明の分野

本発明は、音声認識および自然言語理解の分野に属する。

【背景技術】

【0002】

背景

自動音声認識 (ASR) には、エラーが発生しやすい。

【0003】

ASR は、音声信号に対してスペクトル分析を行い、特徴を抽出する。ASR は、入力音声の音響分析を考慮して、抽出された特徴に対して、複数の音素シーケンスを仮定し、各音素シーケンスは、正しいという可能性を表すスコアを有する。その後、ASR は、各仮定音素シーケンスのスコアを保有しながら、辞書に従って、仮定音素シーケンスを仮定トークンシーケンスにトークン化する。トークンは、英単語などのアルファベット文字、中国語文字などの象形文字、または他の種類の筆記システムの識別可能な要素単位であってもよい。例えば、英語話者 (speaker) がフレーズ「I scream」および「ice cream」をほぼ同様に発音するため、トークン化が曖昧になる。このような曖昧さに対処するために、ASR システムは、話し言葉に共通する隣接ワードトークンの既知頻度の統計を用いて、複数の仮定トークンシーケンスのうち、正しいものを仮定する。例えば、「eat ice cream」のように、単語「eat」の後ろに単語「ice」が続くことはよくあるが、単語「eat」の後ろに単語「I」が続くことは殆どない。したがって、仮定単語シーケンス「I scream」および「ice cream」が単語「eat」の後ろに続く場合、仮定単語シーケンス「ice cream」のスコアが増加するが、仮定単語シーケンス「I scream」のスコアが減少する。

20

30

【0004】

【数 1】

例えば、中国語話者は、フレーズ「交代」(交代) および「胶带」(粘着テープ) を同様に発音する。したがって、音声認識は、話し言葉に共通する隣接トークンの既知頻度の統計を用いて、複数の仮定トークンシーケンスのうち、正しいものを仮定する。例えば、「卷胶带」(粘着テープを巻く) のように、単語「卷」の後ろに単語「胶」が続くことはよくあるが、単語「卷」の後ろに単語「交」が続くことは殆どない。したがって、仮定単語シーケンス「交代」および「胶带」が単語「卷」の後ろに続く場合、仮定単語シーケンス「胶带」のスコアが増加するが、仮定単語シーケンス「交代」のスコアが減少する。

40

【発明の概要】

【発明が解決しようとする課題】

【0005】

従来の音声認識および自然言語理解システムは、比較的不正確で且つ遅いため、文法的に間違っただけの変換を生成し且つリアルタイムの応用に遅延を加える。さらに、従来の音声認識および自然言語理解システムの文法規則の作成および改善が複雑である。また、文法は、通常、ユーザの全ての略式な表現の仕方および近似的な表現の仕方を取り込まないため

50

、十分な適用範囲を有しない。

【課題を解決するための手段】

【0006】

発明の概要

本発明は、正方向および反方向 S L M の組み合わせを用いて音声認識から得られたトークンシーケンスを書き換えるためのシステムをプログラムするための方法、システム、およびソフトウェアに関し、トークンシーケンス内のトークンの確率の組合せに従って、トークンシーケンスのスコアを計算することを含むことができる。

【0007】

本発明は、音声認識の精度および速度、変換文法の正確さおよびリアルタイム変換の遅延に対する技術的な改良並びに自然言語文法の拡張を提供する。

10

【0008】

書換えは、特定のトークンの削除、挿入または置換によるトークンシーケンスの編集を含む。トークンシーケンスにおいて、正方向確率および反方向確率の両方が低いトークンは、削除される対象であるまたは新たなトークンによって置換される対象である。一対のトークンが隣接しており、反方向において前方のトークンが低いスコアを有し且つ正方向において後方のトークンが低いスコアを有する場合、一方のトークンが削除される対象であるまたは両者の間に新たなトークンが挿入される。

【0009】

挿入または置換を行うための新たなトークンを選択することは、正方向 S L M または反方向 S L M のいずれか一方をチェックすることによって確率が最も高いトークンのリストを作成すること、または両方の S L M をチェックすることによって両方の S L M において確率の高いトークンを見出すことを含む。両方の S L M の確率に基づいて新たなトークンを選択する実施形態において、最良の選択は、いずれか一方のリストにおいて最も高い確率を有するトークンではなく、両方のリストにおいてかなり高い確率を有するものである場合が多い。

20

【0010】

削除、挿入、または置換の操作は、新たな組み合わせスコアを有する新たなトークンシーケンスを作成する。いくつかの実施形態は、複数の可能な書換えトークンシーケンスを生成し、各トークンシーケンスのスコアを計算し、最良スコアを有するトークンシーケンスを選択して出力する。いくつかの実施形態は、トークンシーケンスに対して複数の連続書換えを行う。

30

【0011】

いくつかの実施形態は、トークンシーケンスを入力として自然言語パーサに投入し、自然言語パーサは、一連の文法規則に従って、トークンシーケンスの解析スコアを決定する。このような実施形態において、最良のトークンシーケンスの最終的な選択（元のトークンシーケンスまたは任意の特定の書換え）は、文法解析スコアによって決定され、または文法解析スコアとトークンシーケンス確率スコアとの組み合わせによって決定される。一部の文法パーサは、単に、文法に従ってトークンシーケンスが有効であるかまたは無効であることを示す出力を生成する。確率文法のパーサは、解析結果に確率スコアを割り当てる。意味パーサは、意味文法を使用する。意味文法において、解析結果のスコアは、解析結果の意味によって影響される。

40

【0012】

いくつかの実施形態は、例えば、品詞にタグを付けることおよび発話言語において合法的な文法パターンを適用することによって、トークンシーケンスに対して部分的な統語解析を実行する。

【0013】

いくつかの実施形態は、1つ以上の S L M、構文チェック、および1つ以上の人間言語の全てに一般的に使用される表現を含む多様コーパスに対して共通である文法規則を使用する。いくつかの実施形態は、ドメイン固有 S L M、構文チェックおよび文法を使用する

50

ことによって、特定の知識分野に応じて、トークンシーケンスをより正確に書き換えることができる。

【0014】

いくつかの実施形態は、実体のタグ付けを使用することによって、共通クラスを表す特定のタグ付きトークンで既知の特定のトークンを置換する。一例として、一般の都市を表すタグを用いて、都市の名前を表す各トークンを置換する。これによって、あるパターンを満たすトークンシーケンスを数の面で改善し、その結果、統計確率をより精確にすることができる。

【0015】

いくつかの実施形態は、リアルタイム音声の遅延バッファ内で、リアルタイムASRからのトークンシーケンスに対して書換えを行う。いくつかの実施形態は、SLM確率のみを用いて、より短い遅延で書換えを行う。いくつかの実施形態は、より大きな遅延バッファを用いて、完全な表現フレーズの解析を試みる。

10

【0016】

いくつかの実施形態は、最近使用されたトークンの履歴キャッシュを記憶し、履歴キャッシュに存在するトークンの確率ベーススコアを増加させる。このことは、話題に関連する単語が短期間の会話に複数に話される傾向があるという事実を利用して、精度を改善することができる。

【0017】

いくつかの実施形態は、元の入力トークンシーケンスおよび選択された1つの書換えまたは上位のいくつかの書換えの両方をキャッシュに記憶する。これにより、多くの書換えを新たに検索する必要がなく、キャッシュ内の書換えを検索することによって、将来の書換え要求を完了することができる。いくつかの実施形態は、書換えトークンシーケンスおよび当該書換えをもたらした1つ以上の入力トークンシーケンスを記憶する。これにより、有用な情報、特に出現頻度によってソートされた有用な情報を文法開発者に提供することができる。ドメイン固有文法を作成して改善することができる。

20

【0018】

双方向確率的な書換えを使用しない音声認識システムは、比較的高度且つ複雑な文法を必要とするが、双方向確率的な書換えを使用することによって、音声認識システムは、近似表現を文法上正しい表現にマッピングすることによって、比較的簡単な文法の適用範囲を拡張することができる。その結果、それ程優れていない文法開発者でも、双方向確率的な書換えのためのプログラムルールをツールとして採用する音声認識システムに頼り、アプリケーションの要件を満たす特製のドメイン文法を作成することができる。これによって、音声認識システムおよび自然言語理解システムの両方がより高性能でより正確になるという効果をもたらす。

30

【図面の簡単な説明】

【0019】

【図1】一実施形態に従って、音声対応自然言語理解および処理システムを示すフロー図である。

【図2】典型的な英語音声認識エラーを示す図である。

40

【図3】典型的な中国語音声認識エラーを示す図である。

【図4】一実施形態に従って、書換えを示すフロー図である。

【図5A】英語表現を含む簡単なコーパスの例示を示す図である。

【図5B】図5Aのコーパスに従って構築された正方向SLMを示す図である。

【図5C】図5Aのコーパスに従って構築された反方向SLMを示す図である。

【図6A】中国語表現を含む簡単なコーパスの例示を示す図である。

【図6B】図6Aのコーパスに従って構築された正方向SLMを示す図である。

【図6C】図6Aのコーパスに従って構築された反方向SLMを示す図である。

【図7】英語の簡単な自然言語ドメイン文法の例示を示す図である。

【図8】中国語の簡単な自然言語ドメイン文法の例示を示す図である。

50

【図 9】図 7 のドメイン文法によって解析された表現を含む多様コーパスに従って構築された正方向 S L M を示す図である。

【図 10】図 7 のドメイン文法によって解析された表現を含む多様コーパスに従って構築された反方向 S L M を示す図である。

【図 11】一実施形態に従って、タグでトークンを置換した後の書換えを示すフロー図である。

【図 12 A】正しく認識された英語表現の一例を示す図である。

【図 12 B】一実施形態に従って、不正確なトークンエラーによって、置換位置におけるトークンの正方向確率および反方向確率が低くなる、図 12 A の表現の変形例を示す図である。

【図 12 C】一実施形態に従って、余分なトークンエラーによって、挿入位置におけるトークンの正方向確率および反方向確率が低くなる、図 12 A の表現の変形例を示す図である。

【図 12 D】一実施形態に従って、トークン欠落エラーによって、前方のトークンの反方向確率が低くなり、直後のトークンの正方向確率も低くなる、図 12 A の表現の変形例を示す図である。

【図 12 E】一実施形態に従って、トークン重複エラーによって、前方のトークンの反方向確率が低くなり、直後のトークンの正方向確率も低くなる、図 12 A の表現の変形例を示す図である。

【図 13 A】正しく認識された中国語表現の一例を示す図である。

【図 13 B】一実施形態に従って、不正確なトークンエラーによって、置換位置におけるトークンの正方向確率および反方向確率が低くなる、図 13 A の表現の変形例を示す図である。

【図 13 C】一実施形態に従って、余分なトークンエラーによって、挿入位置におけるトークンの正方向確率および反方向確率が低くなる、図 13 A の表現の変形例を示す図である。

【図 13 D】一実施形態に従って、トークンの欠落というエラーによって、前方のトークンの反方向確率が低くなり、直後のトークンの正方向確率も低くなる、図 13 A の表現の変形例を示す図である。

【図 13 E】一実施形態に従って、トークン重複エラーによって、前方のトークンの反方向確率が低くなり、直後のトークンの正方向確率も低くなる、図 13 A の表現の変形例を示す図である。

【図 14】一実施形態に従って、正方向 S L M および反方向 S L M 内の最も高い確率のトークンセットを用いて行った書換えを示すフロー図である。

【図 15 A】一実施形態に従って、英語表現における置換エラーを示す図である。

【図 15 B】一実施形態に従って、2-gram 正方向 S L M を用いて得られた、図 15 A のエラーを置換するための確率が最も高いトークンセットを示す図である。

【図 15 C】一実施形態に従って、2-gram 反方向 S L M を用いて得られた、図 15 A のエラーを置換するための、各方向について最も高い確率に近い共通トークンを有する確率が最も高いトークンセットを示す図である。

【図 16 A】一実施形態に従って、中国語表現における置換エラーを示す図である。

【図 16 B】一実施形態に従って、2-gram 正方向 S L M を用いて得られた、図 16 A のエラーを置換するための確率が最も高いトークンセットを示す図である。

【図 16 C】一実施形態に従って、2-gram 反方向 S L M を用いて得られた、図 16 A のエラーを置換するための、各方向について最も高い確率に近い共通トークンを有する確率が最も高いトークンセットを示す図である。

【図 17 A】一実施形態に従って、英語表現における置換エラーを示す図である。

【図 17 B】一実施形態に従って、3-gram 正方向 S L M を用いて得られた、図 17 A のエラーを置換するための確率が最も高いトークンセットを示す図である。

【図 17 C】一実施形態に従って、3-gram 反方向 S L M を用いて得られた、図 17 A のエ

10

20

30

40

50

ラーを置換するための、各方向について最も高い確率に近い共通トークンを有する確率が最も高いトークンセットを示す図である。

【図 17D】一実施形態に従って、高い確率を有する各 SLM の各共通トークンを用いて誤ったトークンを置換するための書換えスコアの計算を示す図である。

【図 18A】一実施形態に従って、中国語表現における置換エラーを示す図である。

【図 18B】一実施形態に従って、3-gram 正方向 SLM を用いて得られた、図 18A のエラーを置換するための確率が最も高いトークンセットを示す図である。

【図 18C】一実施形態に従って、3-gram 反方向 SLM を用いて得られた、図 18A のエラーを置換するための、各方向について最も高い確率に近い共通トークンを有する確率が最も高いトークンセットを示す図である。

10

【図 18D】一実施形態に従って、各 SLM の高い確率を有する各共通トークンを用いて誤ったトークンを置換するための書換えスコアの計算を示す図である。

【図 19】一実施形態に従って、正方向 SLM および反方向 SLM の最も高い確率のトークンセットを用いて、新たな書換えトークンシーケンスおよび対応するスコアを生成するための書換えを示すフロー図である。

【図 20】一実施形態に従って、正方向 SLM および反方向 SLM の最も高い確率のトークンセットを用いて、多様コーパス SLM に従って、新たな書換えトークンシーケンスおよび対応するスコアを生成するための書換えを示すフロー図である。

【図 21】一実施形態に従って、正方向 SLM および反方向 SLM の最も高い確率のトークンセットを用いて、複数の新たな書換えトークンシーケンスおよび対応するスコアを生成し、最も高いスコアを有する書換えトークンシーケンスを選択するための書換えを示すフロー図である。

20

【図 22】一実施形態に従って、統語解析を用いて、正方向 SLM および反方向 SLM のトークンセットから確率が最も高いトークンの選択を導くための書換えを示すフロー図である。

【図 23】一実施形態に従って、書換えの後に、ドメイン文法を用いて行った自然言語解析を示すフロー図である。

【図 24】一実施形態に従って、異なる英語トークンの挿入に対する文法解析スコアの一例を示す図である。

【図 25】一実施形態に従って、異なる中国語トークンの挿入に対する文法解析スコアの一例を示す図である。

30

【図 26】一実施形態に従って、複数の新たな書換えトークンシーケンスおよび対応するスコアを生成するための書換えを行った後に、ドメイン文法を用いて自然言語解析を行うことによって、スコアが最も高い書換えられたシーケンスを選択することを示すフロー図である。

【図 27】一実施形態に従って、書換えの後に、トークン履歴をキャッシュすることによって、書換えを自動化することを示すフロー図である。

【図 28】一実施形態に従って、入力トークンシーケンスを書換えトークンシーケンスに変換するためのトークン履歴キャッシュの内容を示す図である。

【図 29】一実施形態に従って、書換えトークンシーケンスを入力トークンシーケンスに変換するためのトークン履歴キャッシュの内容を示す図である。

40

【図 30】一実施形態に従って、ユーザとサーバベースの音声対応自然言語理解および処理システムとの対話を示す図である。

【図 31A】一実施形態に従って、非一時的コンピュータ可読媒体であるディスクを示す図である。

【図 31B】一実施形態に従って、非一時的コンピュータ可読媒体であるメモリチップを示す図である。

【図 31C】一実施形態に従って、システムオンチッププロセッサを示す図である。

【図 31D】一実施形態に従って、システムオンチッププロセッサ媒体を示す図である。

【図 32】一実施形態に従って、システムオンチップを示すブロック図である。

50

【図 3 3】一実施形態に従って、サーバシステムを示す図である。

【図 3 4】一実施形態に従って、サーバシステムを示すブロック図である。

【発明を実施するための形態】

【 0 0 2 0 】

詳細な説明

序論

統計言語モデル (S L M) は、表現を含む所定のコーパス内の隣接ワードの統計値を取得する。 S L M を仮定トークンシーケンスに適用することにより、 A S R システムの精度が大幅に改善する。

【 0 0 2 1 】

10

正方向 S L M は、1 つまたは一連の前方トークンが存在する場合に、次のトークンの条件付き確率を表す。反方向 S L M は、1 つまたは一連のの後続トークンが存在する場合に、直前のトークンの条件付き確率を表す。一对の所定のトークンは、正方向および反方向において非常に異なる確率を有することができる。一例として、有名なアメリカ野球選手 Joe DiMaggio を挙げる。反方向 S L M において、家族名 DiMaggio を有する有名人が殆どいないため、DiMaggio に先行する単語が Joe である確率が非常に高い。しかしながら、正方向 S L M において、Joe という名前を有する有名人がたくさんいるため、単語 Joe に単語 DiMaggio が続く確率がかなり低い。

【 0 0 2 2 】

【数 2】

20

また、一例として、中国の三国時代の有名な指揮官「诸葛亮」を挙げる。正方向 S L M において、家族名「诸葛」を有する有名人が殆どいないため、「诸葛」に続く単語「亮」である可能性が非常に高い。しかしながら、反方向 S L M において、「亮」という名前を有する有名人がたくさんいるため、単語「亮」に先行する単語「诸葛」の確率がかなり低い。

表現とは、複数トークンのシーケンスである。様々な実施形態において、トークンは、英語のようなアルファベット文字、または中国語文字などの象形文字である。いくつかの実施形態において、トークンは、単一の概念を表す複数の単語であってもよい。例えば、2 つの単語「New」および「York」または 2 つの文字「紐」および「約」は、共に使用されると、同一の都市を表す。いくつかの実施形態において、異なるトークンは、同様の綴りを有する単語の異なる意味を表す。例えば、名詞の「tire」は、車のタイヤを表し、動詞の「tire」は、疲れることを表す。動詞の「生气」は、怒っていることを表し、名詞の「生气」は、活気または元気を表す。

30

【 0 0 2 3 】

図 1 は、音声対応自然言語処理を行うためのシステムを示す高レベル図である。音声認識モジュール 1 1 は、口頭表現を発したユーザからのスピーチ音声を受け取る。音声認識モジュール 1 1 は、一連のトークンを生成する。いくつかの実施形態は、ユーザが話したことを完璧に理解していないときに、複数の代替的なトークンシーケンスを生成する。このような実施形態の一部は、代替的なトークンシーケンスに確率重みを付ける。

40

【 0 0 2 4 】

いくつかの実施形態は、スピーチ音声以外の方法、例えば、キーボードによるタイピング、タッチスクリーン上のタッピング、目線の移動、腕、手または指によるジェスチャ、または神経電気信号を介して、ユーザからトークンシーケンスを受け取る。

【 0 0 2 5 】

書換えモジュール 1 2 は、入力トークンシーケンスとしてトークンシーケンスを受け取

50

る。書換えモジュール 12 は、後述のプログラムルールをトークンシーケンスに適用することによって、書換えを行う。複数のトークンシーケンスの場合、書換えモジュール 12 は、各々の入力トークンシーケンスを順次にまたは並列に処理する。書換えモジュールは、正方向 S L M 13 および反方向 S L M 14 を用いて、1 つ以上の新たなトークンシーケンスを生成する。いくつかの実施形態は、各新たなトークンシーケンスに対応するスコアを生成する。このスコアは、書換えトークンシーケンスがユーザが表現しようとするものである可能性を示している。いくつかの実施形態は、より良好な書換えトークンシーケンスを生成するように、新たなトークンシーケンスを書換えモジュール 12 に戻して再処理する。これは、例えば、トークンシーケンスに複数のエラーが存在する場合に有用である。

10

【0026】

いくつかの実施形態は、自然言語統語解析モジュール 15 を使用する。統語解析モジュールは、新たなトークンシーケンス、またはいくつかの実施形態では複数の新たなトークンシーケンスを受け取り、ドメイン固有文法 16 に従ってトークンシーケンスを構文解析することによって、自然言語の解釈および解釈に対応する解析スコア。いくつかの実施形態は、複数のドメイン固有文法 16 を使用して、各文法 16 に従って新たなトークンシーケンスを構文解析する。いくつかの実施形態は、書換えモジュール 12 からの書換えスコアを構文解析時に重みとして適用する。いくつかの実施形態は、書換えスコアを用いて、解析スコアに重みを付けることによって、解釈がユーザが意図したものである最良の推定値を生成する。

20

【0027】

いくつかの実施形態は、解析スコアを書換えモジュール 12 に戻し、書換えモジュール 12 は、解析スコアを用いて、多数の可能な書換えから、最も正しいである可能性のある 1 つ以上の書換えを選択する。

【0028】

最終的に、処理モジュール 17 は、最高得点の解釈を受け取り、その意図を判断し、意図の判断に応じて解釈を処理することによって、最も可能な有用な結果を生成する。処理モジュール 17 は、結果を生成し、当該結果をシステム出力としてユーザに提供する。

【0029】

書換えモジュール 12 を使用することによって、実施形態は、ユーザの表現の正しいトークンシーケンスを決定し、したがって正しい解釈および有用な結果を生成する可能性が高くなる。さらに、書換えスコアのフィードバックおよび解析スコアのフィードバックを使用することによって、実施形態は、ユーザの表現の正しいトークンシーケンスを決定し、したがって正しい解釈および有用な結果を生成する可能性がさらに高くなる。その結果、このような実施形態は、誤って認識された入力トークンシーケンスまたは文法的に不正確な入力トークンシーケンスを文法的に正しいものに変換することによって、文法の言語適用範囲（すなわち、文法によって認識されるコーパスの部分）を増加させ、これによって、文法は、不備のある変形を効率的にサポートすることができ、エラー訂正を効率的に行うことができる。

30

【0030】

自然言語の双方向確率的な書換えおよび選択は、文法が解釈できる表現の範囲を広げる、即ち、一種の効率的な文法規則の圧縮である。これにより、小さいサイズの文法を用いて、より大きいサイズの文法を必要とする表現の範囲をサポートすることができ、したがって、同一サイズの文法を用いて、より広い範囲の人間表現を解釈することができる。

40

【0031】

また、人間言語の自動機械翻訳を行うためのシステムは、改善された翻訳精度を得ることができる。このような実施形態の一部は、書換えを翻訳モジュールに入力する前に、書換えを A S R 結果に適用する。書換えは、翻訳する必要がある表現の範囲を減らすことによって、より簡単な翻訳ルールで、必要な翻訳精度を達成することができる。

【0032】

50

図30は、サーバ実装システムに含まれた図1の要素を示している。このシステムは、携帯電話アプリまたは家庭内デバイスとして実装された典型的な仮想アシスタントのように、ネットワークを介してサーバに接続されたクライアントを介して、ユーザから表現を受け取る。いくつかの実施形態は、典型的な自動車システム、高セキュリティシステムおよび極小電力システムのように、自己完結型であり、ネットワークから独立して動作する。

【0033】

図2は、ユーザが天気に関連するクエリを生成しようとし、天気に関連する文法によって正しく解析できなかった英単語のトークンシーケンスの種類および例を示している。

【0034】

図3は、ユーザが天気に関連するクエリを生成しようとし、天気に関連する文法によって正しく解析できなかった中国語文字のトークンシーケンスの種類および例を示している。

【0035】

図2および図3に示された各例示的なシーケンスは、4種類のエラー、即ち、不正確なトークン、余分なトークン、欠落トークン、および重複トークンのうち、1つ以上のエラーによって引き起こされる。各種類のエラーは、3種類の編集、即ち、トークンの挿入、トークンの削除、およびトークンの置換のうち、いずれか1つによって訂正することができる。いくつかの実施形態は、トークン置換という種類の編集を有していないが、トークン削除およびトークン挿入の組み合わせで同様の効果を達成する。トークン挿入およびトークン置換は各々、入力トークンシーケンスに存在しなかった新たなトークンを作成する。以下に説明するように、プログラムルールシステムは、先行トークンおよび後続トークンを考慮して、トークンの条件付き確率に基づいて、トークンの書換えを行うための編集の種類および位置を決定する。

【0036】

トークン置換の場合、置換位置は、低い確率を有するトークンにある。トークン挿入の場合、挿入位置は、2つのトークンの間にあり、前方のトークンが低い反方向確率を有し、後方のトークンが低い正方向確率を有する。トークン削除の場合、削除位置は、双方向に低い確率を有するトークンにある、または低い反方向確率を有する前方のトークンおよび低い正方向確率を有する後方のトークンのうちいずれか一方のトークンにある。

【0037】

いくつかの実施形態は、単一の書換えにおいて複数のトークンの置換を行う。いくつかの実施形態は、特定の書換えにおいて単一のトークンの置換を行う。このような実施形態の一部は、複数の書換えを繰り返して行う。

【0038】

図4は、トークン置換を行うためのトークンシーケンス書換えフローを示している。正方向トークン確率モジュール41は、入力トークンシーケンスを受け取り、正方向SLM42を用いて、入力トークンシーケンス内の各トークンの正方向確率を生成する。反方向確率モジュール43は、入力トークンシーケンスを受信し、反方向SLM44を用いて、入力トークンシーケンス内の各トークンの反方向確率を生成する。編集モジュール45は、各トークンの正方向確率および反方向確率を受け取り、正方向および反方向両方の確率が低いトークンを見出し、そのトークンを別のものに置換することによって、新たな書換えトークンシーケンスを生成する。図4に示すように、いくつかの実施形態は、少なくとも1つのトークンが十分に低い確率を有する場合に限り、条件付きのトークン置換を行う。

【0039】

いくつかの実施形態は、図4のように動作するが、編集モジュール45において、置換を行う代わりに、低い確率を有するトークンの削除を行う。

【0040】

いくつかの実施形態は、図4のように動作するが、編集モジュール45において、置換

10

20

30

40

50

を行う代わりに、反方向に低い確率を有する前方のトークンおよび正方向に低い確率を有する後方のトークンを含む一対の隣接トークンを特定する。このような実施形態の一部は、一対のトークンの間に新たなトークンを挿入する。このような実施形態の一部は、一対のトークンのうち、第1トークンまたは第2トークンを置換する。

【0041】

閾値

いくつかの実施形態は、確率を閾値と比較することによって、低い確率を決定する。いくつかの実施形態は、閾値を使用せず、単に最も低い確率の位置を書き換える。いくつかの実施形態において、正方向の閾値と反方向の閾値とは異なる。いくつかの実施形態において、閾値は、隣接するトークンの確率に依存する。

10

【0042】

良好な閾値レベルは、コーパス内の全てのトークンの確率分布において、トークンの大部分を通過させることができる点である。閾値が低くなると、より積極的な書換えが行われる。積極的な書換えは、より深刻なトークンエラーを修正することができるが、入力トークンシーケンスを話者の意図を誤って表すものに書き換える危険性がある。

【0043】

トークンシーケンスの先頭のトークンの正方向確率は、シーケンスの先頭から数個後のトークンの正方向確率よりも遥かに広く分散されている。同様に、トークンシーケンスの末尾のトークンの反方向確率は、シーケンスの末尾から数個先のトークンの反方向確率よりも遥かに広く分散されている。いくつかの実施形態は、シーケンス内の第1のトークンまたは第1のトークン群に対して、シーケンス中の他のトークンと異なる閾値を使用する。

20

【0044】

いくつかの実施形態において、閾値は、ユーザプロファイルの特性に依存する。具体的には、各ユーザは、自分が作成した有効な表現を含むコーパスから構築された個人SLMを持っている。このような実施形態は、共通SLMに対して適切な閾値とユーザ固有SLMに応じた閾値との組み合わせである閾値を使用する。これによって、単純な表現をするユーザに対して、最高精度の改良を与え、複雑な表現をする傾向のあるユーザに対して、確率の低いトークンの書換えを減らすことができ、便利である。

【0045】

いくつかの実施形態において、閾値を使用せず、正方向、反方向、または正方向および反方向の両方の組み合わせにおいて、確率の最も低いトークンを選択して置換する。

30

【0046】

SLMの構築

最も有用な書換えの選択は、精確なSLMの有無に依存する。SLMは、書き換える表現と同一種類の表現を含むコーパスから構築された場合、最も精確である。例えば、Twitter（登録商標）のツイートの表現を含むコーパスは、ニューヨークタイムズ紙の新聞記事の表現を含むコーパスとは非常に異なるSLM確率を有する。同様に、仮想アシスタント用の表現を含むコーパスは、一般的に天気ドメインに特有の表現を含むコーパスとは異なるSLM確率を有する。

40

【0047】

精度を高めるために、SLMを構築するために使用されるコーパスは、正しい表現、即ち、適切な自然言語文法が解析できる表現または確率的な文法によって解析されるときに高いスコアを有する表現のみを含むべきである。さらに、異なる特定の实体（場所名など）を参照する表現を含むコーパスを用いてSLMを構築すると、異なる特定の实体に各々対応する低確率シーケンスを多く作成する。汎用タグを用いて特定の实体のトークンを置換することによって得られたSLMは、タグを用いて实体を置換することによって得られた同様の表現に対して遥かに高い確率を有する。タグ置換によって構築されたSLMを利用するためには、トークン書換えシステムがトークンシーケンスの書換えを行う前に、対応するタグ置換を行う必要がある。

50

【 0 0 4 8 】

【 数 3 】

図 5 A および図 6 A は、例示としての小さなコーパスを示している。各コーパスは、5 つの表現を含み、各表現は、2 つのトークンを含む。図 5 B および図 6 B は、コーパスから構築された正方向 S L M を示している。各行は、一連の直近トークンに対応する（1 つのみのトークンを含むシーケンスをサポートする）。各列は、シーケンスの後続トークンに対応する。各セルは、確率を示す。例えば、図 5 B の S L M において、トークン「who」の後に、次のトークンが「is」である確率は、 $1/2$ である。図 6 B の S L M において、トークン「是」の後に、次のトークンが「哪里」である確率は、 $1/4$ である。図 5 C および図 6 C は、コーパスから構築された反方向 S L M を示している。各行は、一連の後方トークンに対応する（1 つのみのトークンを含むシーケンスをサポートする）。各列は、シーケンスの前方トークンに対応する。例えば、図 5 C の S L M において、トークン「is」の場合、前方トークンが「who」である確率は、 $1/4$ である。図 6 C の S L M において、トークン「哪里」の場合、次のトークンが「是」である確率は、 $1/2$ である。なお、行列として表現された正方向および反方向 S L M は、互いの転置ではない。

10

20

【 0 0 4 9 】

図 7 は、天気ドメインにおける英語表現を解釈するための簡単な自然言語文法の一例を示している。図 8 は、天気ドメインにおける中国語表現を解釈するための簡単な自然言語文法の一例を示している。これらの文法は、単なる例示にすぎない。これらの文法は、あまりにも少ないフレーズを定義しているため、天気に関するクエリを表現する多くの態様をサポートすることができない。

【 0 0 5 0 】

この文法は、タグ <date> および <place> を用いて、日付および地名を表す。実体にタグを付けることができる適切な自然言語処理システムは、これらの実体を検出し、特定の日付または地名が表現に存在する場合に、文法規則を適用する。

30

【 0 0 5 1 】

図 9 は、図 7 の文法におけるフレーズに各々一致する多数の表現を含むコーパスから構築された正方向 S L M を示している。各行は、一連の直近トークンに対応する（1 つのトークンのみを含むシーケンスをサポートする）。他の実施形態は、2 つ以上のトークンのシーケンスを用いて、次のトークンをより精確に予測する。各列は、シーケンスの後続トークンに対応する。各セルは、確率を示す。記号 <s> および </s> は、トークンシーケンスの先頭および末尾を示す。例えば、シーケンスの先頭に位置する最初の単語が「what」である確率は、0.69 であり、シーケンスの末尾に位置する最後の単語が日付である確率は、0.67 である。単語「what」が単語「the」に続く確率は、0.01 であるため、人々が「the what」を言うのは滅多にないことを示している。

40

【 0 0 5 2 】

図 10 は、図 7 の文法におけるフレーズに各々一致する多数の表現を含むコーパスから構築された反方向 S L M を示している。各行は、一連の後続トークンに対応する（1 つのみのトークンを含むシーケンスをサポートする）。各列は、一連の先行トークンに対応する。例えば、「weather」という単語で終わるトークンシーケンスの確率は、0.128 である。「what」という単語の後ろに、単語「will」が続く確率は、0.51 であるため、人々が「what will」を言うのは比較的一般的であることを示している。

【 0 0 5 3 】

図 5、6、9 および 10 の例は、2-gram (bigram ととも言う) S L M を示している。即ち、これらの S L M は、1 つの特定のトークンに隣接する各トークンの確率を示している。

50

一部の S L M は、3-gram を使用する。すなわち、このような S L M は、2 つの隣接トークンからなる特定のシーケンスに隣接する各トークンの確率を示している。単語「will」の後ろに単語「be」が続く確率と 2 つの単語「what will」の後ろに単語「be」が続く確率とは、異なる。いくつかの S L M は、4-gram またはより高次の n-gram を使用する。

【 0 0 5 4 】

いくつかの実施形態は、多種多様なトピックに関連する表現を含む多様コーパスから S L M を構築する。これによって、多種多様な表現、例えば、特定の間言語の全ての可能な表現をサポートすることができ、便利である。

【 0 0 5 5 】

いくつかの実施形態は、1 つのみのドメインに関連する表現を含むコーパスから S L M を構築する。このようなコーパスは、表現の多様性がより少ないため、ドメイン内の表現に対する n-gram 間の確率の差を高める。その結果、ドメインに対して正しい可能性が最も低いトークンが、より精確に選択され書換えられる。多様な表現から構築されたコーパスは、編集するための最良のトークンをあまり精確に選択していないため、特定のドメインと同様の精度を達成するために、より多くの書換えを選択する必要がある。

【 0 0 5 6 】

いくつかの実施形態は、多数の表現をフィルタリングすることによって、ドメイン固有文法が有効であると解釈できる表現またはドメイン固有確率文法が閾値を超えるスコアで解析できる表現を生成する。このような実施形態は、合格した表現のみを S L M 訓練コーパスに使用する。このような実施形態の一部は、コーパスをさらにフィルタリングすることによって、1 つ以上の他の文法においてより高い解釈スコアを有しない表現のみを生成する。いくつかの実施形態は、解釈を行う時に、会話状態履歴、位置およびユーザプロフィールなどの他のコンテキスト情報、またはその両方を考慮する。

【 0 0 5 7 】

表現のコーパスがなくても、一組の文法規則が解析できる一組の表現を用いて、S L M を構築することが可能である。確率文法を表現生成器として使用することによって、このような S L M を生成することができる。この場合、文法規則の重みは、トークンシーケンスの確率を決定する。文法のみから S L M を構築することは、文法設計者が人々によって行われた文法ドメインに関連する表現を全て予測し且つ文法の重みを推測することができる限り、成功する。実際には、S L M 確率は、表現のコーパスから構築された場合により精確である。文法から S L M を構築することは、より精確な S L M を構築するために使用できる多数の表現を収集するためのシステムをブートストラップするときに、よく行われる。文法のみから生成された S L M 確率と、適切なコーパスから構築することによって決定された確率とを組み合わせることも可能である。

【 0 0 5 8 】

タグ付け

図 1 1 は、トークン置換を行うためのトークンシーケンス書換えフロー図を示している。このフロー図は、タグモジュール 1 1 6 が入力トークンシーケンスを受け取り、タグを用いて、トークンを置換することによって、確率モジュール 4 1 および 4 3 用のトークンシーケンスを作成することを除き、図 4 のフロー図と同様である。

【 0 0 5 9 】

「タグ」とは、トークンのクラスに対して共通トークンである。例えば、<city> タグは、トークン「London」および「Beijing」並びにトークンシーケンス「New」「York」のいずれかを置換することができる。<time> タグは、トークンシーケンス「1:23」、「1:23 pm」、「01:23 pm」、「13:23」、および「one twenty-three in the afternoon」（午後 1 時 2 3 分）のいずれかを置換することができる。

【 0 0 6 0 】

いくつかの実施形態は、ドメイン固有タグを使用する。例えば、天気ドメインは、<weather condition> タグを有することができる。<weather condition> タグは、「sunny」（晴れ）、「cloudy」（曇り）、「foggy」（霧）、「rainy」（雨）、「windy」（風

10

20

30

40

50

）、「hot」（熱い）、および「cold」（寒い）などのトークンを置換することができる。この実施形態は、まず、不正確なフレーズ「it sunny or rainy will be」からなるトークンシーケンスを受け取り、次に、タグ<weather condition>の各特定のインスタンスと元のトークンとの関連性を保存すると共に、タグインスタンス<weather condition 1>を用いてトークン「sunny」を置換し、タグインスタンス<weather condition 2>を用いてトークン「rainy」を置換することによって、タグを付けたトークンシーケンスを「will it be <weather condition 1> or <weather condition 2>」に書換え、最後に、元のトークンを用いて<weather condition>タグの各インスタンスを置換することによって、書換えトークンシーケンス「will it be sunny or rainy」（天気は晴れですかまたは雨ですか）を生成する。天気ドメイン文法は、入力トークンシーケンス「it sunny or rainy will be」を解析することができないが、書換えトークンシーケンス「will it be sunny or rainy」を解析することができる。

10

【0061】

いくつかの実施形態において、ドメイン固有文法を使用する場合、文法に存在していない全てのトークンに、<unknown>（未知）タグを付ける。この実施形態において、タグを用いて未知トークンを書換え、書換えた後に、元のトークンを用いて未知トークンのタグを置換する。一実施形態において、SMSテキストメッセージドメインの文法には、「送信」トークン、「テキスト」トークンおよび「メッセージ」トークンを含む。当該実施形態において、「メッセージ牛馬および鶏をテキストメッセージとして送信する」というトークンシーケンスを受信すると、このトークンシーケンスを「メッセージ<unknown 1> <unknown 2> および <unknown 3> をテキストメッセージとして送信する」のようにタグを付ける。当該実施形態は、「<unknown 1> <unknown 2> および <unknown 3> をテキストメッセージとして送信する」のように書換えを行い、元のトークンを用いてタグを置換することによって、書換えトークンシーケンス「牛馬および鶏をテキストメッセージとして送信する」を生成する。英語表現の一般的な用法によって、実施形態は、「<unknown>を送信する」というトークンシーケンスをよく受け取り、「メッセージ<unknown>」というトークンシーケンスを滅多に受け取らないため、最初に出現したトークン「メッセージ」に対して削除編集を実行する。

20

【0062】

いくつかの実施形態は、多様コーパスから構築されたSLMを使用する。このような実施形態の一部は、コーパスに低頻度で現れる全ての単語を<低頻度>タグを付ける。全ての低頻度単語にタグを付けることによって、高頻度単語の確率をより精確にすることができる。

30

【0063】

いくつかの実施形態は、コーパス内の品詞の検出およびトークンのラベル付けプロセスを行う。異なる品詞として機能するトークンの場合、例えば、「fast」というトークンが形容詞（速い）、動詞（断食する）および名詞（断食期間）として使用する場合、実施形態は、SLMにおいて、異なる品詞として使用されているトークンに応じて異なるエントリーを含む。このような実施形態において、書換えられるトークンシーケンスにおけるトークンの確率を計算する前に、トークンシーケンスに対して部分的な統語解析を行うことによって、シーケンス内の各トークンの可能な品詞を識別する。これによって、SLMにおいてラベル付けられたトークンの確率がより精確になる。

40

【0064】

いくつかの実施形態は、コーパス内の表現および入力トークンシーケンスの両方に対して、同義語置換を行う。例えば、このような実施形態は、トークン「fast」を用いて、トークン「rapid」、「quick」および「speedy」のいずれかを置換する。その結果、多様な表現を含む所定のコーパスに対するSLM統計がより精確になる。

【0065】

いくつかの実施形態は、コーパス内の表現および入力トークンシーケンスの両方に対して、同音異義語置換を行う。例えば、このような実施形態は、単一のトークンを用いて、

50

英語トークン「to」、「too」および「two」を置換する。

【0066】

いくつかのドメイン固有の実施形態は、書換えられない保留セットのドメイン固有トークンを有する。例えば、SMSメッセージのドメインにおいて、実施形態は、トークン「テキスト」、「メッセージ」および「送信」を置換、削除または編集操作として挿入できないトークンとして含む。一部のドメイン文法フォーマットは、保留キーワードを定義することができる。

【0067】

確率からエラー種類を検出する

正方向および反方向の各々において低い確率を有するトークンの位置のコロケーションまたは隣接は、可能なエラーの種類（不正確なトークン、余分なトークン、欠落トークン、および重複トークン）に関するヒントを与え、その結果、トークンシーケンスを書き換えるための可能な最良種類の編集（トークンの挿入、トークンの削除、およびトークンの置換）に関するヒントを与える。

【0068】

図12Aは、反方向の確率および正方向の確率の両方が共に高い（H）である表現に対応する一連の英単語トークンを示している。図13Aは、反方向の確率および正方向の確率の両方が共に高い（大）である表現に対応する一連の中国語文字トークンを示している。

【0069】

【数4】

図12Bおよび図13Bは、英単語および中国語文字のトークンエラー（例えば、図12Bの「in」ではなく「pin」、図13Bの「的」の代わりに「跑的」）を含む表現を各々示している。反方向において、不正確なトークンに続く全てのトークンが高い確率を有し、正方向において、不正確なトークンに先行する全てのトークンが高い確率を有する。反方向および正方向の各々において、不正確なトークンが低い確率を有する。

【0070】

図12Cおよび図13Cは、英単語および中国語文字の余分なトークンエラーを含む表現を各々示している。反方向において、余分なトークンに続く全てのトークンが高い確率を有し、正方向において、余分なトークンに先行する全てのトークンが高い確率を有する。反方向および正方向の各々において、余分なトークンが低い確率を有する。

【0071】

【数5】

図12Dおよび図13Dは、英単語および中国語文字のトークン欠落エラー（図12Dにおいて「weather」と「Hawaii」との間にトークン「in」が欠落しており、図13Dにおいて「的」と「怎么样」との間に「天气」欠落している）を含む表現を各々示している。反方向において、欠落トークンの直前のトークンが低い確率を有するが、欠落トークンに続く全てのトークンが高い確率を有する。正方向において、欠落トークンの前の全てのトークンが高い確率を有するが、欠落トークンの直後のトークンが低い確率を有する。

【0072】

図12Eおよび図13Eは、英語および中国語文字のトークン重複エラーを含む表現を示している。反方向において、重複トークンの第1のトークンが低い確率を有するが、第2のトークンおよび全ての後続トークンが高い確率を有する。正方向において、重複トークンの第1のトークンを含み、第1のトークンの前の全てのトークンが高い確率を有するが、重複トークンの第2のトークンが低い確率を有する。

【0073】

10

20

30

40

50

従って、反方向および正方向の両方において低い確率を有するトークンが疑わしいトークンであり、削除または置換編集が適切である可能性の高い場所を示している。前方のトークンが反方向に低い確率を有し、後方のトークンが正方向に低い確率を有する一対の隣接する同一ではないトークンは、疑わしいトークン対であり、挿入編集が適切である可能性が高い場所を示している。前方のトークンが反方向に低い確率を有し、後方のトークンが正方向に低い確率を有する一対の隣接する同一のトークンは、疑わしいトークン対であり、同一のトークンのうち、1つのトークンの削除が適切である可能性の高い場所を示している。

【0074】

当然なことに、音声認識は、確率的なことである。トークンシーケンスは、話し言葉の音素を含む1つ以上の仮定シーケンスから得られ、これらの音素は、デジタル処理した音響信号サンプルのフレームから得られる。いくつかの実施形態は、仮定音素シーケンス内の各音素の認識精度の信頼度を表すスコアを生成する。いくつかの実施形態において、サンプルのフレームを処理することによって、仮定音素のスコアに寄与する値を生成する。いくつかの実施形態において、音素シーケンスをトークンシーケンスにトークン化する際に、少なくとも部分的にトークン内の音素の音響スコアに基づいて、各トークンのスコアを計算する。

【0075】

いくつかの実施形態は、トークンの音響スコアによって重み付けられた反方向SLM確率および正方向SLM確率に少なくとも部分的に基づいて、正方向確率および反方向確率スコアを計算する。その結果、一実施形態において、得られた品質の悪いスピーチ音声、例えば、一部のトークンの間に大音量の背景音を有するスピーチ音声の場合、これらのトークンを集中的に書き換えることができる。

【0076】

書換えの採点および比較

スコアは、整数、浮動小数点数、または英数字などの様々な方法で表すことができる。いくつかの実施形態において、スコアは、有効性のブール値（有効／無効）として表される。

【0077】

様々な実施形態は、反方向確率が閾値未満である位置、または正方向確率が閾値未満である位置、または反方向確率と正方向確率との組み合わせが閾値未満である位置を特定することによって、書換えを作成する。このような実施形態の一部は、まず、閾値未満の確率スコアを有する位置で単一の編集を各々行うことによって、複数の代替的な書換えを生成し、その後、第1セットの書換えトークンシーケンス内の第2の位置で更なる編集を行うことによって、更なる書換えを行う。これによって、書換えツリーまたは書換え格子を作成することができ、置換Aの後ろに置換Bが続くことと、（別の場所で）置換Bの後ろに置換Aが続くこととは同様の結果を有するということを説明することができる。一実施形態において、閾値が高いまたはスピーチ音声の品質が悪い場合、書換えの数が使用可能なシステムリソースを大きく超える可能性がある。このような制限に対応するために、様々な実施形態において、幅優先探索アルゴリズム、深さ優先探索アルゴリズム、または最良優先アルゴリズムを用いて、例えば、最良の書換えトークンシーケンスのスコアなどの有用性測定値に基づいて、書換えツリーまたは書換え格子を構築することによって、最良の書換えを見出す。

【0078】

上述したように、様々な実施形態において、反方向の確率が最も低い位置、正方向の確率が最も低い位置、または同一のトークンまたは隣接するトークンのいずれかの反方向確率および正方向確率の組み合わせが最も低い位置で、編集を行うことによって、書換えを作成する。様々な実施形態において、組み合わせは、積、合計または同等物などの関数である。このような実施形態の一部において、複数の書換えを作成することによって一組の仮定書換えを生成し、各仮定書換書の書換えスコアを計算することによって、最良な（最

10

20

30

40

50

も可能な正しい)書換えを選択することができる。いくつかの実施形態において、最も低い確率を有する位置および2番目低い確率を有する位置で各々編集を行い、得られた第1対の書換えのトークンシーケンスを編集することによって、複数の書換えを作成する。

【0079】

図14は、トークン置換を行うためのトークンシーケンスの書換えフロー図を示している。このフロー図は、トークン置換モジュール145が正方向SLMおよび反方向SLMの各々から確率が最も高いトークンのリストを読み取る点を除き、図4のフロー図と同様である。各位置の正方向確率が最も高いトークンのリストは、前方のN個トークンに対して、確率が最も高いトークンのセットおよび対応する確率を含む。各位置の反方向確率が最も高いトークンのリストは、後方のN個トークンに対して、確率が最も高いトークンのセットおよび対応する確率を含む。挿入または置換するためのトークンを選択するために、トークン置換モジュール145は、両方のリスト(リストの交点)に存在するトークンを検索し、各リスト内のトークンの確率に基づいて新たなトークンを選択する。

【0080】

いくつかの実施形態において、トークンが両方のリストに存在しない場合、SLMから得られた次の確率が最も高い単語をリストに追加する。様々な実施形態において、トークンが両方のリストに存在しない場合、正方向リストから確率が最も高いトークンを選択する、または反方向リストから確率が最も高いトークンを選択する、またはトークンシーケンスの書換えをスキップする。

【0081】

図15Aおよび図16Aは各々、3つのトークンを含み、中央のトークンが不正確である英単語および中国語文字のトークンシーケンスを示している。図15Bおよび図16Bは、5つのトークンおよび(1つの先行トークンから各トークンの確率を計算する)2-gram正方向SLMから得られた各トークンの確率を含むリストを示している。図15Cおよび図16Cは、5つのトークンおよび2-gram反方向SLMから得られた各トークンの確率を含むリストを示している。これらのリストは、2つの共通トークンを有する。第1の共通トークンは、各リストにおいて2番目に高い確率を有する。第2の共通トークンは、正方向リストにおいて5番目に高い確率を有し、反方向リストにおいて3番目に高い確率を有する。第1の共通トークンは、両方のリストにおいてより一般的であるため、トークン置換編集の最良選択である。

【0082】

図17Aおよび図18Aは各々、5つのトークンを含み、中央のトークンが不正確な英単語および中国語文字のトークンシーケンスを示している。図17Bおよび図18Bは、5グラムのリストと、5つのトークンおよび(2つの先行トークンから各トークンの確率を計算する)3-gram正方向SLMから得られた各トークンの確率を含むリストを示している。図17Cおよび図18Cは、5つのトークンおよび3-gram反方向SLMから得られた各トークンの確率を含むリストを示している。正方向SLMと反方向SLMとは、同様の長さのn-gramを使用する必要がない。これらのリストは、2つの共通トークンを共有する。第1の共通トークンは、正方向リストにおいて3番目に高い確率を有し、反方向リストにおいて2番目に高い確率を有する。第2の共通トークンは、正方向リストにおいて5番目に高い確率を有し、反方向リストにおいて1番目に高い確率を有する。2つのトークンが両方のリストに共通であるが、共通トークンの確率の順番が両方のリストにおいて異なるため、どちらがトークン置換編集の最良選択であるかを計算する必要がある。

【0083】

いくつかの実施形態において、各リスト内の共通トークンの確率を加算し、確率の加算値が最大であるトークンを選択する。いくつかの実施形態において、各リスト内の共通トークンの確率を乗算し、確率の乗算値が最大であるトークンを選択する。いくつかの実施形態において、各リスト内の共通トークンの確率の対数値を加算し、確率の対数値の和が最大であるトークンを選択する。これによって、絶対値が異なっているが、乗算と同様の相対順位を与えることができる。

【 0 0 8 4 】

図 1 7 D および図 1 8 D は、加算による各共通トークンの確率の計算、乗算による各共通トークンの確率の計算、および対数値の加算による各共通トークンの確率の計算を示している。加算は、第 2 のトークンのより高い書換えスコアを見出すが、乗算および対数値の加算は、第 1 のトークンのより高い書換えスコアを見出す。いくつかの実施形態において、乗算および対数値の加算が好ましい。確率の乗算および対数値の加算は、両方向でより合理的なトークンにとって有利であるが、確率の加算は、一方の方向では合理的であり、他方の方向では不合理なトークンにとって有利である。

【 0 0 8 5 】

いくつかの所定のスピーチを翻訳するために、いくつかの同音異義語のうち、正しいものを検出することは、困難な課題である。いくつかの実施形態において、入力トークンシーケンス内の疑わしいトークンと同音異義語であるトークンの確率に好ましい重みを適用することにより、書換えスコアを計算する。その結果、このような実施形態において、周りのトークンに照らして最も可能な同音異義語を遥かに正確に選択することができる。

10

【 0 0 8 6 】

トークンは、最近の会話に使用された場合、より高い確率を有する。いくつかの実施形態は、最近使用されたトークンをキャッシュする。このような実施形態の一部は、正帰還ループによって誤った翻訳を誘起する可能性を最小化するために、高い音響スコアを有するトークンのみをキャッシュする。ある仲間によって使用されたトークンは、他の仲間によって使用される確率が高いため、いくつかの実施形態は、最近の会話に複数の仲間によって使用されたトークンをキャッシュする。

20

【 0 0 8 7 】

図 1 9 は、トークン置換を行うためのトークンシーケンスの書換えフロー図を示している。このフロー図は、トークン置換モジュール 1 9 5 が書換えトークンシーケンスおよび対応する書換えスコアの両方を生成する点を除き、図 1 4 のフロー図と同様である。いくつかの実施形態において、書換えスコアは、各トークンの正方向確率と各トークンの反方向確率との積である。いくつかの実施形態において、書換えスコアは、音響スコアによって重み付けられた各トークンの正方向確率と音響スコアによって重み付けられた各トークンの反方向確率との積である。いくつかの実施形態において、実体値を置換できるように実体のタグ付けを使用しない場合、特定の人間の位置または存在などの環境パラメータ値に対する特定の実体の近接度に従って、書換えスコアに重みを付ける。

30

【 0 0 8 8 】

図 2 0 は、トークン置換を行うためのトークンシーケンスの書換えフロー図を示している。このフロー図は、トークン置換モジュール 2 0 5 が多様コーパス S L M 2 0 7 からトークンの条件付き確率を読み取る点を除き、図 1 9 のフロー図と同様である。正方向 S L M 4 2 および反方向 S L M 4 4 は、1 つの会話ドメインに特有の表現から構築される。多様コーパス S L M 2 0 7 は、多くのドメインに亘る表現から構築され、例えば、これらのドメインが単一の言語を使用するという事実のみによって制限された多様なソースの表現から構築される。いくつかの実施形態は、多様コーパス S L M 2 0 7 内の各トークンに対応する条件付き確率に基づいて、入力トークンシーケンス内の各トークンのドメイン固有条件付き確率をスケーリングする。いくつかの実施形態は、多様コーパス S L M 2 0 7 内の各トークンに対応する条件付き確率に基づいて、最も可能なリスト内の各置換トークンのドメイン固有条件付き確率をスケーリングする。いくつかの実施形態は、単一方向の多様コーパス S L M を使用し、いくつかの実施形態は、各方向の多様コーパス S L M を使用する。

40

【 0 0 8 9 】

図 2 1 は、トークン置換を行うためのトークンシーケンスの書換えフロー図を示している。このフロー図は、トークン置換モジュール 2 1 5 が複数の書換えトークンシーケンスと、書換えられた各トークンシーケンスに対応する書換えスコアとを生成する点を除き、図 1 9 のフロー図と同様である。選択モジュール 2 1 8 は、書換えスコアを用いて、複数

50

の代替的な書換えから、最良の書換えトークンシーケンスを選択する。一部の代替的な書換えは、同一の場所で異なる編集が行われる。一部の代替的な書換えは、異なる場所で編集が行われる。

【 0 0 9 0 】

解析スコアを計算するための解析

いくつかの実施形態は、S L M確率および重み付きスコアよりもさらに進化した処理を用いて、最良の書換えを決定する。図 2 2 は、統語解析を使用する実施形態を示している。示された実施形態は、トークン置換モジュール 2 2 5 が複数の書換えを生成し、最高のスコアを有する 1 つの書換えを選択して、新たな書換えトークンシーケンスを生成する点を除き、図 1 4 と同様である。また、トークン置換モジュール 2 2 5 は、統語解析入力を用いて、書換えスコアに重みを付ける。

10

【 0 0 9 1 】

統語解析モジュール (syntactic analysis module) 2 2 9 は、入力トークンシーケンスを受け取り、品詞として使用された各トークおよび品詞として使用された各トークンの頻度を決定する。トークン置換モジュール 2 2 5 は、異なる品詞としてのトークンの配置の合法性に従って、構文規則を適用する。トークン置換モジュール 2 2 5 は、隣接するトークンの文脈において構文上合法的なトークンのみに限定して、新たなトークンを選択する。各書換えに対して、トークン置換モジュール 2 2 5 は、特定の品詞としてのトークンの配列の合法性に従って、書換えトークンシーケンスにスコアを付ける。

20

【 0 0 9 2 】

いくつかの実施形態は、自然言語パーサを用いて、文法に従って新たな書換えトークンシーケンスを解析することによって、構文解析スコアを生成する。図 2 3 は、トークン置換を行うためのトークンシーケンスの書換えフロー図を示している。このフロー図は、図 4 のフロー図と同様である。しかしながら、自然言語パーサ 2 3 6 は、トークン置換モジュール 4 5 から出力された新たな書換えトークンシーケンスを受け取る。自然言語パーサは、自然言語文法 2 3 7 によって定義された文法規則に従って、新たな書換えトークンシーケンスを解析する。

【 0 0 9 3 】

いくつかの実施形態は、ブール値スコアを生成する簡単な文法を使用する。すなわち、文法は、トークンシーケンスを解析することができるか否かに関係ない。いくつかの実施形態は、整数スコアまたは浮動小数点スコアをパーサに与える確率文法を使用する。

30

【 0 0 9 4 】

いくつかの実施形態において、文法規則は、異なる言い回しの可能性を定義する。文法は、2 つの異なる表現を解析することによって同様の意味を得る (即ち、同様の解釈データ構造を生成する) ことができ、2 つの異なる表現のうち、より一般的に使用される表現を示す指標または当該表現内により一般的に使用されるフレーズまたは単語を示す指標を含むことができる。いくつかの実施形態は、話者の間に表現の言い回しおよびその言い回しを使用する可能性を考慮して、表現の解析スコアを計算する。このような実施形態の一部は、文法規則内のフレーズまたは単語の重みを乗算することによって、解析スコアを計算する。いくつかの実施形態は、統語解析スコアと意味スコアとを組み合わせることによって、解釈のスコアを決定する。

40

【 0 0 9 5 】

【数 6】

図 2 4 は、天気ドメインの確率文法に従って同様の意味を有する 2 つの英語表現を示している。一人の話者が単語「for」を使用し、もう一人の話者が単語「in」を使用する。より多くの話者が単語「in」を含むフレーズを好むため、そのフレーズのスコアが比較的高くなる。図 2 5 は、天気ドメインの確率文法に従って、同様の意味を持つ 2 つの中国語表現を示している。一部の話者が単語「里」を使用し、一部の話者が単語「儿」を使用する。より多くの話者が単語「里」を含むフレーズを好むため、そのフレーズのスコアが比較的高くなる。

10

【0 0 9 6】

文法を有するいくつかの実施形態は、最近の会話状態の履歴情報をキャッシュする。例えば、会話状態は、自然言語の文法規則内の実体スロットの意味値を含むことができる。「ニューヨーク」は、「X の天気は何ですか」（X は、場所である）という表現に応答する天気関連文法規則の「場所」スロットに存在する実体種類「場所」の意味値である。したがって、会話状態を記憶する実施形態は、「ニューヨークの天気は何ですか」という表現の後に、最も最近に参照した場所「ニューヨーク」の記録を保存するであろう。

【0 0 9 7】

また、会話状態は、過去に質問した問題の種類を含むこともできる。また、会話状態は、実体の関連パラメータの識別子を含むこともできる。さらに、会話状態は、限定詞に関する情報および情報フィルタリング情報を含むことができる。いくつかの実施形態は、解析スコアとして表すことができる会話状態に照らして、トークンの意味的な顕著性に従って、トークンの解析スコアに重みを付ける。

20

【0 0 9 8】

いくつかの実施形態は、複数のドメイン固有文法に従って、書換えトークンシーケンスを解析しようと試みる。少なくとも 1 つのドメイン文法に従って、許容スコアで書換えトークンシーケンスを解析した場合、書換えは成功である。

【0 0 9 9】

図 2 6 は、トークン置換を行うためのトークンシーケンスの書換えフロー図を示している。このフロー図は、自然言語パーサ 2 6 6 が、一組の書換えトークンシーケンスおよび関連スコアを受け取り、文法 2 3 7 に従って各トークンシーケンスを解析することによって解析スコアを生成し、解析スコアと書換えスコアとの最適な組み合わせを有する書換えトークンシーケンスを出力として選択する点を除き、図 1 9 のフロー図と同様である。このような実施形態の一部は、複数の文法に従って解析を行う。したがって、これらの実施形態は、事実上、音響スコア、S L M による書換えスコアおよび文法解析スコアの組み合わせを有する出力トークンシーケンスを生成する。

30

【0 1 0 0】

書換えセットの反復作成

いくつかの実施形態は、成功した書換えを見出すと、処理を完了し、成功した書換えトークンシーケンスを出力として提供する。いくつかの実施形態は、入力トークンシーケンスの可能な書換えのリストを生成する。これらの可能な書換えは、同一の疑わしい場所で異なる編集を行うことによって、異なる疑わしい場所で編集を行うことによって、またはその両方によって得ることができる。

40

【0 1 0 1】

いくつかの実施形態において、複数の間違っただ単語または複数の余分な単語、欠落単語、または重複単語を引き起こしたエラーを修正するために、書換えフローの追加反復を通過するように書換えトークンシーケンスを供給する。反復は、書換えスコア、文法解析スコア、またはその両方の組み合わせに基づいて、書換えの簡単なリストを作成することによって、または再処理するための書換えを選択することによって、継続することができる

50

。選択は、深さ優先探索ツリーアルゴリズム、幅優先探索ツリーアルゴリズム、または最高スコア優先探索ツリーアルゴリズムを試行することによって行われる。

【0102】

図1は、新たなトークンシーケンスを反復的に書き換えるプロセスを示している。いくつかの実施形態において、書換え段階12の後に、新たなトークンシーケンスの書換えスコアに従って、これらのトークンシーケンスを再処理する。いくつかの実施形態において、構文解析段階14の後に、新たな書換えトークンシーケンスの構文解析スコアに従って、これらの書換えトークンシーケンスを再処理する。

【0103】

連続翻訳

トークンシーケンスの書換えは、連続翻訳応用、例えば、テレビライブ放送の自動字幕付けまたは書き取りに有用であり得る。

【0104】

双方向確率的な書換えは、反方向SLM確率を計算するエンドポイントがある場合に有用である。これは、十分に高いトークン化仮定スコアを有するシーケンスのうち、直近のN個(Nは、反方向SLM確率の計算に使用されるトークンの数である)トークンを利用可能にするトークンバッファによって可能になる。この方法は、翻訳時に少なくともN個トークンの時間に相当する遅延を必要とする。

【0105】

英語話者は、1分間に150トークンで話す傾向がある。中国語話者は、1分間に約300文字で話す傾向がある。多くの中国語トークンが2文字または3文字で構成されているため、話す時のトークン速度も毎分約150トークン、即ち毎秒2.5トークンである。したがって、2-gram反方向SLMを使用する双方向書換えは、少なくとも1秒のバッファ遅延を必要とする。

【0106】

連続音声応用の場合、文法解析を双方向確率的な書換えに適用するためには、完全な表現に含まれる全てのトークンまたは少なくとも完全なフレーズが、シーケンスを解析できると判断するために必要である。このような応用において、書換えおよび選択の前に、フレーズ境界検出アルゴリズムを行う必要がある。当該技術分野において、様々なフレーズ境界検出アルゴリズムが既知である。一部のフレーズ境界検出アルゴリズムは、文法に基づいたものである。一部のフレーズ境界検出アルゴリズムは、話すときの声の変化を検出することによって機能する。一部のフレーズ境界検出アルゴリズムは、対応するビデオシーンの変化を検出することによって機能する。また、十分な長さの休止は、コミットメントを強要することができる。

【0107】

人間は最大25トークンを含む表現を話す傾向があるため、1分間に150トークンのトークン速度で話す場合、最大10秒のバッファおよび潜在的な遅延が必要である。これは、映画または事前録画されたテレビ番組には許容できない字幕遅延であるが、テレビのニュースフィードまたは一部のスポーツ放送には合理的である。一度に完全な表現を処理することによって、字幕表示は、テレビライブ放送の場合に一般的である従来の単語ごとの表示ではなく、映画および事前録画されたテレビ番組の場合に一般的であるように、表現ごとのテキストを追加する。

【0108】

いくつかの実施形態は、リアルタイム放送のビデオストリームを遅延させることにより、表現の分離並びに文法に基づいた採点および選択による書換えに必要な遅延と同調させることができる。

【0109】

いくつかの実施形態は、音素仮定、トークン化仮定、およびフレーズ境界仮定などの仮定を用いて、音声翻訳プロセスを実行する。このような実施形態は、仮定ツリーを構築し、当該ツリーの各枝は、スコアを有する。このような実施形態の一部は、様々なトークン

10

20

30

40

50

化された表現仮定の各々に対して双方向確率的な自然言語書換えを実行し、得られた書換えの解析スコアを計算し、各枝の最良書換えの解析スコアに従って、ツリーの最良枝を選択する。

【0110】

履歴キャッシングおよび逆キャッシング

いくつかの実施形態は、1つ以上のキャッシング技術を用いて性能を向上させる。いくつかの実施形態は、少なくとも限定された期間において、高い音響仮定スコアを有するトークンをキャッシュする。このような実施形態は、正方向SLMおよび反方向SLMの最も高い確率のトークンリストからトークンを選択する場合、キャッシュに存在するトークンの確率スコアを増加させる。この方法は、会話中の単語の時間的局所性を利用して、認識精度を向上させる。様々な実施形態は、実体にタグを付ける前に単語をキャッシュすることによって実体名の認識を向上させる、または実体にタグを付けた後に単語をキャッシュすることによって、実体に依存しない書換の統合を容易にする。

10

【0111】

いくつかの実施形態は、入力トークンシーケンスと新たな書換えトークンシーケンスとの間のマッピングをキャッシュする。図27は、トークンシーケンスの書換えフロー図を示している。このフロー図は、図4のフロー図と同様である。但し、シーケンス検査モジュール278は、入力トークンシーケンスを受け取る。その後、このシーケンス検査モジュールは、入力トークンシーケンスと新たな書換えトークンシーケンスとの間のマッピングを含むキャッシュ279からの読み込みを進める。シーケンス検査モジュール278は、キャッシュから入力トークンシーケンスを見付けた場合、関連する新たな書換えトークンシーケンスをトークン置換モジュール275に提供する。トークン置換モジュール275は、シーケンス検査モジュール278から新たな書換えトークンシーケンスを受け取ると、新たな書換えトークンシーケンスを直接に出力する。そうでない場合、トークン置換モジュール275は、高い確率SLMトークンのリストを使用した書換えを含む適切な書換え解析を実行する。したがって、このフロー図は、過去に受け取ったトークンシーケンスの最良書換えを得るための処理集約ツリー検索 (processing-intensive tree search) を実行する必要がない。

20

【0112】

図28は、トークン書換えキャッシュの7つのエントリを示している。(タグ付き実体を含む)各入力トークンシーケンスは、ちょうど1つの新たな書換えトークンシーケンスに関連付けられる。

30

【0113】

いくつかの実施形態は、新たな書換えトークンシーケンスを複数の入力トークンシーケンスと関連付けることができる逆キャッシュを使用する。図29は、逆トークン書換えキャッシュの5つのエントリを示している。最後のキャッシュエントリは、5つの異なる入力トークンシーケンスの各々から作成された共通の書換えである。キャッシュ内の最後の新たなトークンシーケンスを作成した各入力トークンシーケンスの頻度は、括弧に示されている。いくつかの実施形態は、キャッシュの内容を分析することによって、特定の新たな書換えトークンシーケンスに書換えられた最も頻度の高い入力トークンシーケンスを見出す。これは、文法の開発者が、文法を改善するために、共通に使用されている代替フレーズを見出し、これらを文法規則に追加するまたは入力トークンシーケンスが表現の意図と一致しない書換を扱うための異なる文法規則を作成する時に有用である。例えば、「will the weather be good」(天気が良いでしょうか)を「what will the weather be」(天気がどうなるでしょうか)に書き換える場合、システムエンジニアは、前者のフレーズをサポートする新しい文法規則を作成し、予報された気象条件に基づいて、「はい」または「いいえ」で応答することができる。

40

【0114】

いくつかの実施形態は、文法自動帰納モジュールを含み、この文法自動帰納モジュールは、逆書換えキャッシュを用いて、新しい文法規則を推論することによってまたは既存の

50

文法規則を拡張することによって、適用範囲を改善する。

【0115】

特定のトークンの重みを調整することまたは規則を変更することによって確率文法を改善することにより、システムオペレータが正確なSLMを生成する能力を高める。さらに、確率的な文法を改善することにより、正確なドメイン固有SLMを構築するためのコーパスのフィルタリングを改善する。

【0116】

その他の考慮事項

本明細書に示され説明された実施形態は、多くのオプション特徴の例示である。示され説明された特徴の異なる組み合わせによって、多くの実施形態が可能である。本明細書に示され説明された例は、英語および中国語を使用している。様々な実施形態は、他の言語または言語の組み合わせに対しても同様の効果を奏する。本明細書に示され説明された例は、天気ドメインを使用している。様々な実施形態は、他のドメインまたはドメインの組み合わせに対しても同様の効果を奏する。本明細書に示され説明された例は、トークンを置換するためのモジュールを示している。同様に配置されたモジュールは、トークンの削除およびトークンの挿入を行うことができる。

【0117】

図30は、クライアント-サーバ型アーキテクチャを使用する仮想アシスタントの一実施形態を示している。ユーザ301は、スマートフォン302に表現を話す。スマートフォン302は、仮想アシスタントアプリを実行している場合、サーバベースの仮想アシスタント用のリモートクライアント装置として動作する。スマートフォン302は、仮想アシスタント用のクエリとして、ユーザ301からの表現を送信する。クライアント装置302は、コンピュータネットワーク303を介してクエリ音声を通信用し、表現音声をサーバ304に送信する。サーバ304は、コンピュータプロセッサを用いて、非一時的なコンピュータ可読媒体に格納されたコードを実行する。コンピュータプロセッサは、命令を処理することにより、表現音声に対して自動音声認識305を実行して、仮定トークンシーケンスを生成する。コンピュータプロセッサは、仮定トークンシーケンスを入力として使用して、書換え306を実行することによって、新たな仮定の書換えトークンシーケンスを生成する。コンピュータプロセッサは、自然言語処理(NLP)307を用いて、新たな仮定の書換えトークンシーケンスを解析する。書換えは、NLP解析307の結果を用いて、別の書換えを行うかまたはサーバ304からの新たな書換えトークンシーケンスを出力するかを判断する。許容可能な高いスコアを有する解析を選択すると、コンピュータプロセッサは、選択された解釈を処理すること(308)によって、所望の結果を生成する。この処理は、例えば、サーバ304内でローカルに実行されてもよく、第三者システムのサービスを用いて実行されてもよい。例えば、「天気は何ですか」というクエリを受けた場合、処理工程308は、携帯電話の現在位置を含むアプリケーションプログラミングインターフェイス(API)リクエストを第三者の天気サービスに送信する。第三者の天気サービスは、気象条件のテキスト記述を用いて、APIリクエストに回答する。仮想アシスタントは、他の多くの情報源および多く種類のサービスを提供することができる。サーバ304は、APIリクエストに対するテキスト応答などの処理結果を用いて、ユーザ301に与える音声応答のテキストを生成する。コンピュータは、テキスト音声変換(TTS)309を用いて、音声応答のテキストを処理することによって、ユーザに与える音声応答を生成する。サーバ304は、ネットワーク303を介して、音声をクライアント302に送信し、クライアント302は、スピーカを介して、この音声をユーザに出力する。他の実施形態において、多くの他の変形が可能である。例えば、他の実施形態において、サーバ304は、音声応答の代わりにまたは音声応答に加えて、テキスト応答をクライアント302に提供する。

【0118】

図31Aは、いくつかの実施形態に従って、非一時的なコンピュータ可読媒体である回転ディスク311を示している。

【 0 1 1 9 】

図 3 1 B は、いくつかの実施形態に従って、非一時的なコンピュータ可読媒体であるフラッシュメモリチップ 3 1 2 を示している。

【 0 1 2 0 】

図 3 1 C は、プリント回路基板に半田付けられるボールグリッドアレイを備えたパッケージ型システムオンチップ装置 3 1 3 の底面を示す。システムオンチップ 3 1 3 は、いくつかの実施形態に係るプロセッサである。

【 0 1 2 1 】

図 3 1 D は、システムオンチップ装置 3 1 3 の上面を示している。

図 3 2 は、いくつかの実施形態に従って、システムオンチップチップ 3 1 3 内の機能要素を示すブロック図である。コンピュータプロセッサコア 3 2 1 およびグラフィックプロセッサコア 3 2 2 は、相互接続 3 2 3 および R A M インターフェイス 3 2 4 を介して一時データを記憶するためのランダムアクセスメモリ (R A M) と通信し、相互接続 3 2 3 およびフラッシュインターフェイス 3 2 5 を介してフラッシュ R A M と通信し、並びに相互接続 3 2 3 およびネットワークインターフェイス 3 2 8 を介してサーバと通信することによって処理を行う。表示インターフェイス 3 2 6 は、表示情報の利用をユーザに提供し、I / O インターフェイス 3 2 7 は、ユーザ入力を受け取り、ユーザに出力を提供する。

10

【 0 1 2 2 】

図 3 3 は、いくつかの実施形態に従って、ラック型マルチプロセッササーバ 3 0 3 を示している。

20

【 0 1 2 3 】

図 3 4 は、いくつかの実施形態に従って、サーバ 3 0 4 内の機能要素のブロック図を示している。コンピュータプロセッサ 3 4 1 およびグラフィックプロセッサ 3 4 2 は、相互接続 3 4 3 を介して一時データを記憶するための R A M 3 4 4 と通信し、ネットワークインターフェイス 3 4 5 を介して他のサーバおよび装置と通信することによって処理を行う。

【図 1】

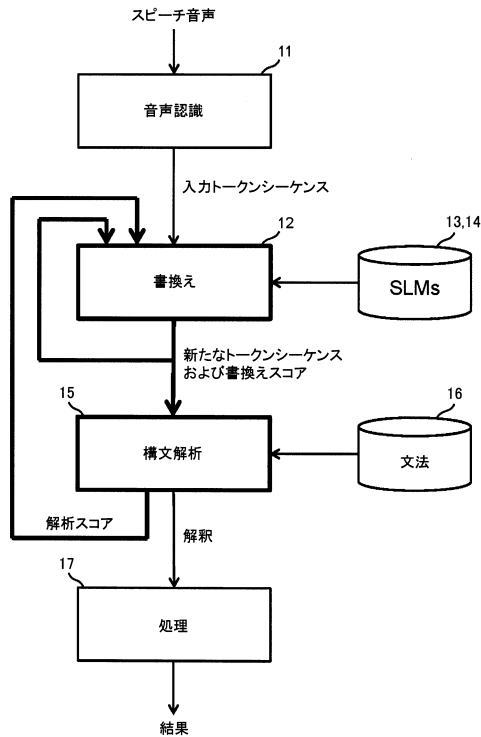


FIG. 1

【図 2】

不正確な単語
will there be andy rain today

不要な単語
will there be like rain today
will there the be rain today

欠落
is it going to ____
____ going to rain today

吃音／重複
what what is the weather
what how is the weather

拡張／冗語
do we have rain for today

背景音声
it is working is it going to rain today

未対応言い回し
today will it rain

FIG. 2

【図 3】

不正確な文字
今天活下雨吗

不要な文字
今天会去下雨吗
今天的会下雨吗

欠落
今天会下——
——雨吗

吃音／重複
今天天气怎么怎么样
今天天气什么怎么样

拡張／冗語
我们今天是不是要会有雨

背景音声
他会不会今天会下雨吗

未対応言い回し
下雨今天会吗

FIG. 3

【図 4】

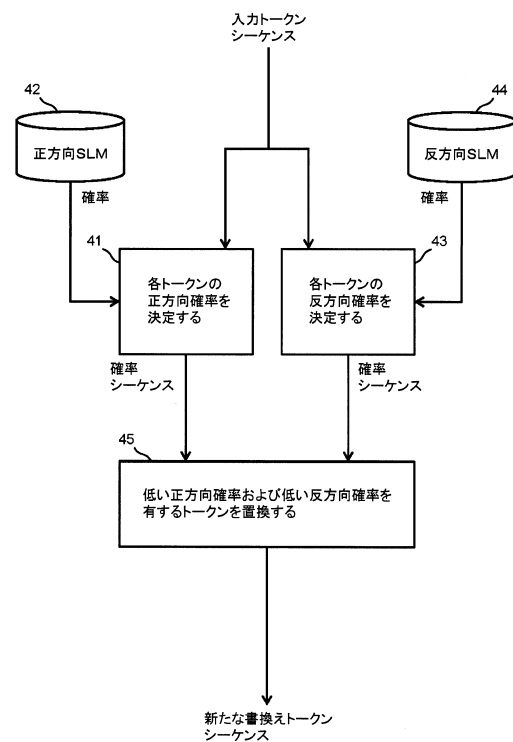


FIG. 4

【図 5 A】

コーパス
 what is
 who is
 when is
 what is
 who did

FIG. 5A

【図 5 B】

正方向SLM確率

	is	did
who	½	½
what	1	0
when	1	0

FIG. 5B

【図 5 C】

反方向SLM確率

	who	what	when
is	¼	½	¼
did	1	0	0

FIG. 5C

【図 6 A】

コーパス
 是什么
 是谁
 是哪里
 是什么
 去哪里

FIG. 6A

【図 6 B】

正方向SLM確率

	哪里	什么	谁
是	¼	½	¼
去	1	0	0

FIG. 6B

【図 6 C】

反方向SLM確率

	是	去
哪里	½	½
什么	1	0
谁	1	0

FIG. 6C

【図 7】

受け入れる文法

what is the weather
 what will the weather be on <date>
 what is the weather going to be <date>
 what is the weather in <place>
 what is the weather at <place>

FIG. 7

【図 8】

受け入れる文法

天气怎么样
 <日期>的天气如何
 <日期>会是什么天气
 <地点>的天气如何
 在<地点>是什么天气

FIG. 8

【図 9】

	<s>	at	<place>	in	to	going	<date>	on	be	will	weather	the	is	what	</s>
<s>	0.69	0.02	0.018	0.02	0.02	0.019	0.019	0.02	0.02	0.02	0.041	0.02	0.02	0.02	0.03
at	0.02	0.02	0.022	0.02	0.02	0.023	0.023	0.03	0.02	0.19	0.022	0.02	0.02	0.02	0.04
<place>	0.018	0.022	0.017	0.02	0.02	0.018	0.018	0.02	0.02	0.02	0.017	0.02	0.02	0.02	0.03
in	0.02	0.02	0.017	0.02	0.02	0.018	0.018	0.02	0.02	0.02	0.017	0.02	0.02	0.02	0.03
to	0.02	0.02	0.017	0.02	0.02	0.018	0.018	0.02	0.02	0.02	0.017	0.02	0.02	0.02	0.03
going	0.02	0.02	0.017	0.02	0.02	0.018	0.018	0.02	0.02	0.02	0.017	0.02	0.02	0.02	0.03
<date>	0.019	0.023	0.017	0.02	0.02	0.018	0.018	0.02	0.02	0.02	0.017	0.02	0.02	0.02	0.03
on	0.02	0.03	0.02	0.02	0.02	0.018	0.018	0.02	0.02	0.02	0.017	0.02	0.02	0.02	0.03
be	0.02	0.02	0.02	0.02	0.02	0.018	0.018	0.02	0.02	0.02	0.017	0.02	0.02	0.02	0.03
will	0.02	0.02	0.02	0.02	0.02	0.018	0.018	0.02	0.02	0.02	0.017	0.02	0.02	0.02	0.03
weather	0.02	0.02	0.02	0.02	0.02	0.018	0.018	0.02	0.02	0.02	0.017	0.02	0.02	0.02	0.03
the	0.02	0.02	0.02	0.02	0.02	0.018	0.018	0.02	0.02	0.02	0.017	0.02	0.02	0.02	0.03
is	0.02	0.02	0.02	0.02	0.02	0.018	0.018	0.02	0.02	0.02	0.017	0.02	0.02	0.02	0.03
what	0.02	0.02	0.02	0.02	0.02	0.018	0.018	0.02	0.02	0.02	0.017	0.02	0.02	0.02	0.03
</s>	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03

FIG. 9

【図 14】

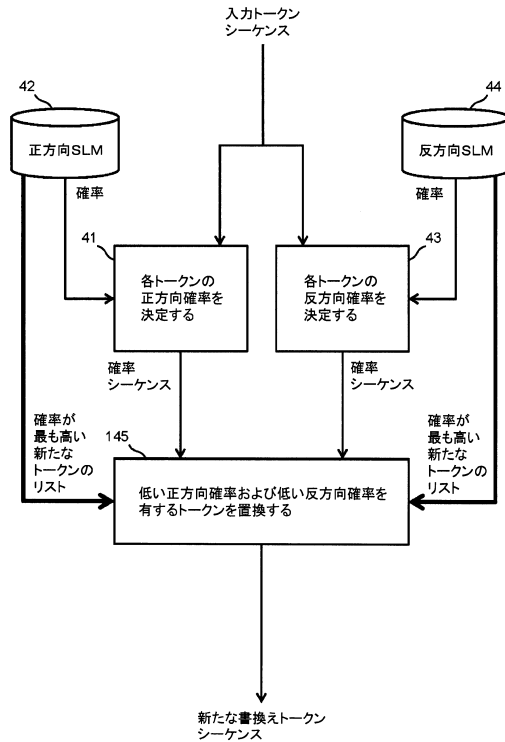


FIG. 14

【図 15 A】

weather pin Hawaii
FIG. 15A

【図 16 C】

— 天気
查询 .072
的 .059
和 .036
多云 .012
恶劣 .003
FIG. 16C

【図 17 A】

the weather pin Hawaii today
FIG. 17A

【図 17 B】

the weather
today .083
forecast .076
for .052
now .038
in .029
FIG. 17B

【図 17 C】

Hawaii today
in .102
for .071
to .054
over .032
and .011
FIG. 17C

【図 15 B】

weather
forecast .073
in .069
Channel .033
data .013
and .005
FIG. 15B

【図 15 C】

Hawaii
to .072
in .059
and .036
Honolulu .012
love .003
FIG. 15C

【図 16 A】

夏威夷 天气
FIG. 16A

【図 16 B】

夏威夷
岛的 .073
的 .069
旅游 .033
海滩 .013
和 .005
FIG. 16B

【図 17 D】

書換えスコア
in .029 + .102 = 0.131
.029 * .102 = 0.0030
 $\log(.029) + \log(.102) = -2.53$
for .052 + .071 = 0.123
.052 * .071 = .0037
 $\log(.042) + \log(.071) = -2.43$
FIG. 17D

【図 18 A】

今天 天气 夏威夷的天气
FIG. 18A

【図 18 B】

今天—
是 .083
运势 .076
在 .052
日期 .038
去 .029
FIG. 18B

【図 18 C】

夏威夷
去 .102
在 .071
来 .054
感受 .032
漫步 .011
FIG. 18C

【図 18 D】

書換えスコア

去 $.029 + .102 = 0.131$
 $.029 * .102 = 0.0030$
 $\log(.029) + \log(.102) = -2.53$

在 $.052 + .071 = 0.123$
 $.052 * .071 = .0037$
 $\log(.042) + \log(.071) = -2.43$

FIG. 18D

【図 19】

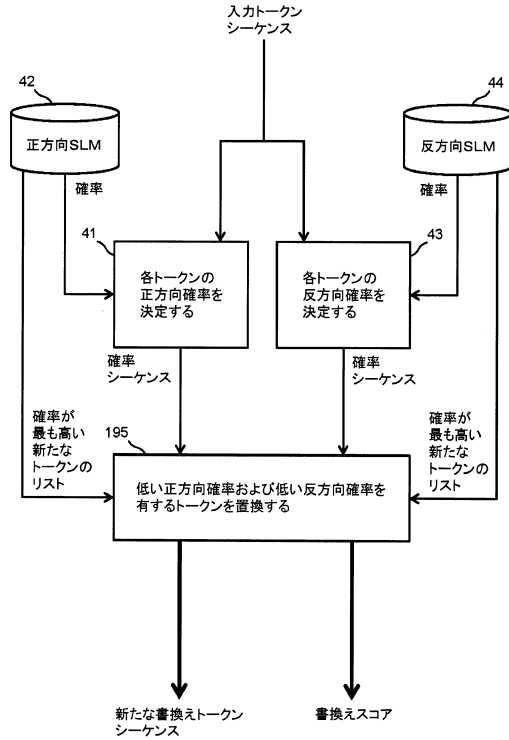


FIG. 19

【図 20】

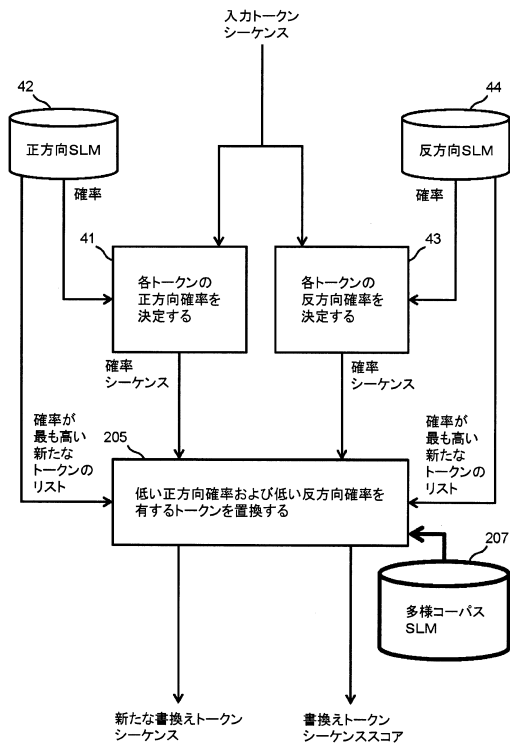


FIG. 20

【図 21】

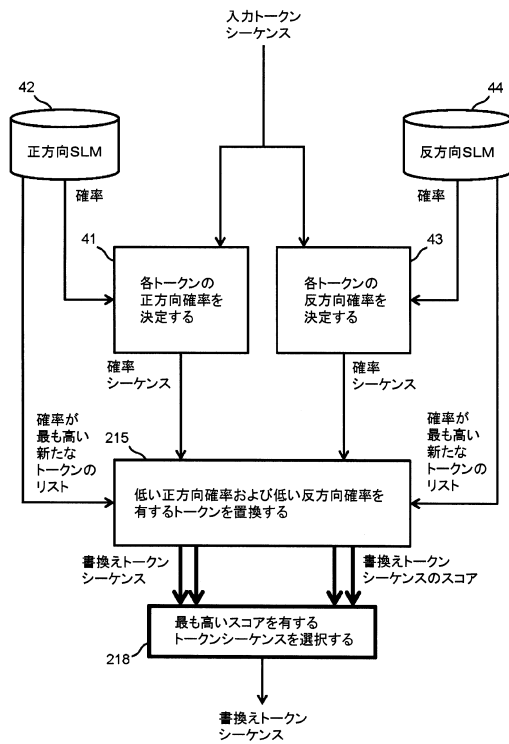


FIG. 21

【図 22】

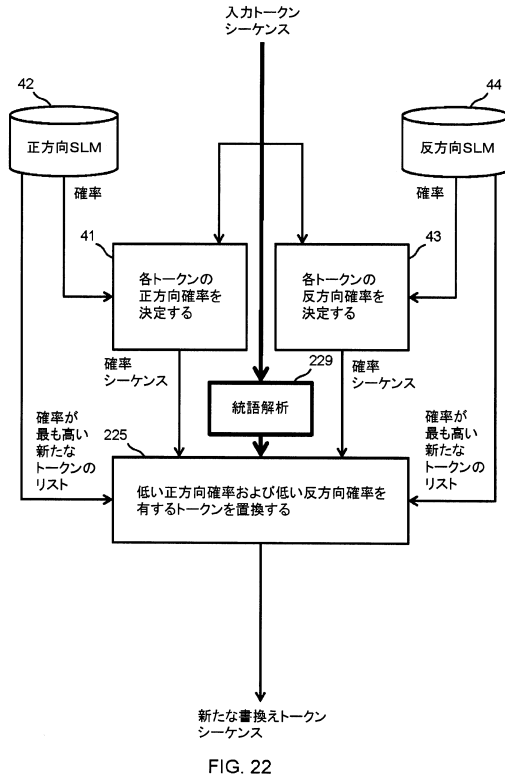


FIG. 22

【図 23】

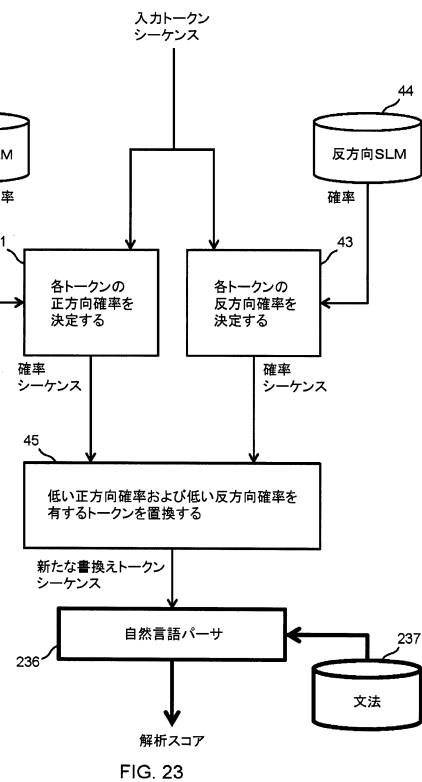


FIG. 23

【図 24】

what's the weather for Hawaii today
解析スコア : .29

what's the weather in Hawaii today
解析スコア : .35

FIG. 24

【図 25】

夏威夷在哪儿?
解析スコア : .29

夏威夷在哪里?
解析スコア : .35

FIG. 25

【図 26】

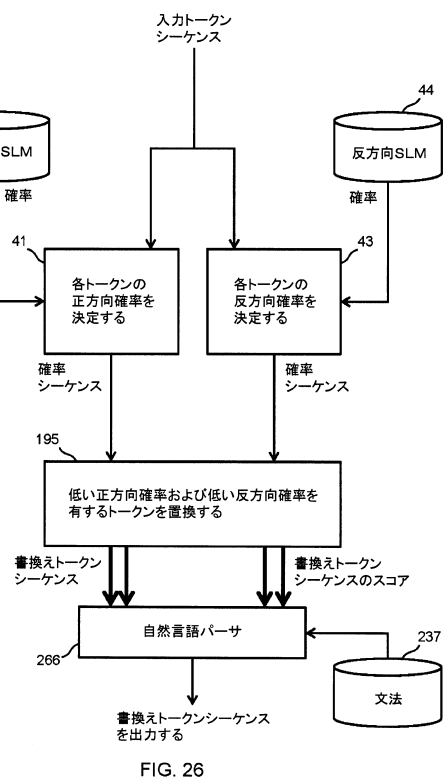


FIG. 26

【図 27】

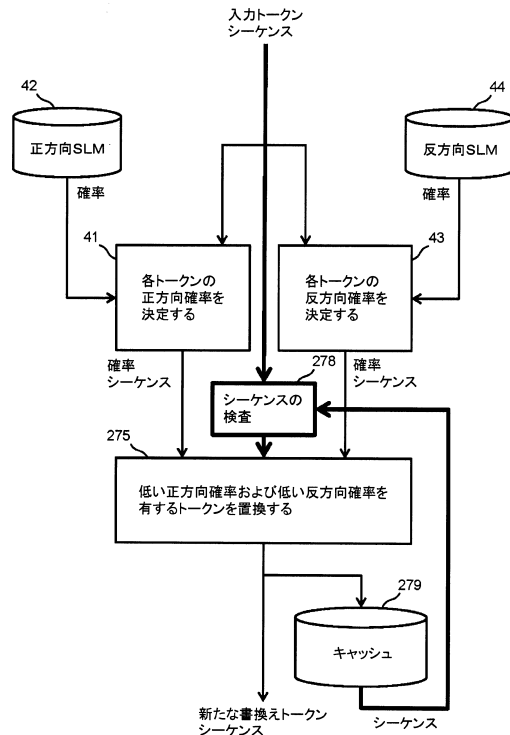


FIG. 27

【図 28】

入力 : what the weather pin <place>
<date>
書換え: what's the weather in <place>
<date>

入力 : will there be andy rain <date>
書換え: will there be any rain <date>

入力 : is it going to
書換え: is it going to rain <date>

入力 : what the weather pin <place>
<date>
書換え: what's the weather in <place>
<date>

入力 : do we have rain for <date>
書換え: will there be rain on <date>

入力 : what the weather in <place>
<date>
書換え: what's the weather in <place>
<date>

入力 : <date> will it rain
書換え: will it rain <date>

FIG. 28

【図 29】

書換え: will there be any rain <date>
入力 : will there be andy rain <date>

書換え: is it going to rain <date>
入力 : is it going to

書換え: will there be rain on <date>
入力 : do we have rain for <date>

書換え: will it rain <date>
入力 : <date> will it rain

書換え: what's the weather in <place>
<date>
入力 : what's the weather for <place>
<date> (3)
入力 : what's the weather on <place>
<date> (2)
入力 : what the weather pin <place>
<date>
入力 : what the weather in <place>
<date>
入力 : what the weather for <place>
<date>

FIG. 29

【図 30】

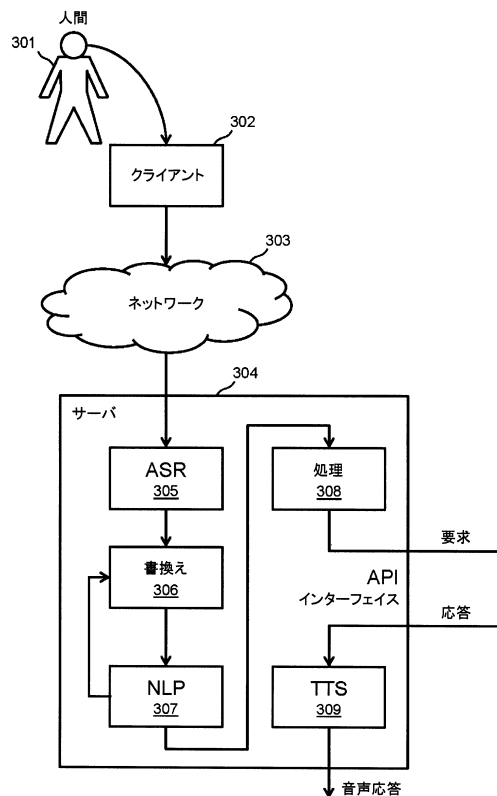


FIG. 30

【図 3 1 A】

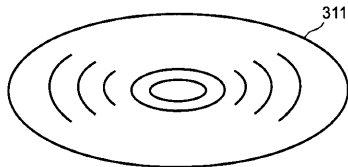


FIG. 31A

【図 3 1 D】

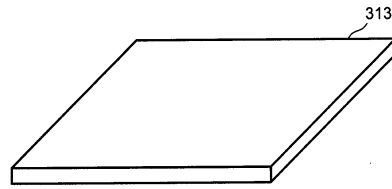


FIG. 31D

【図 3 1 B】

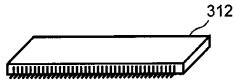


FIG. 31B

【図 3 1 C】

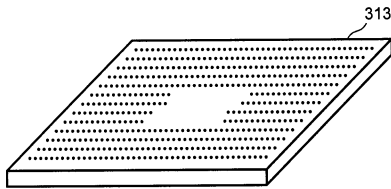


FIG. 31C

【図 3 2】

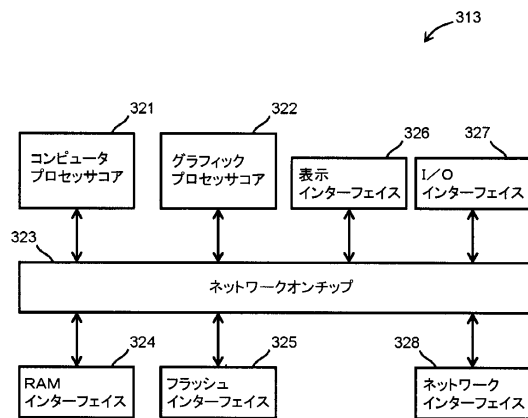


FIG. 32

【図 3 3】

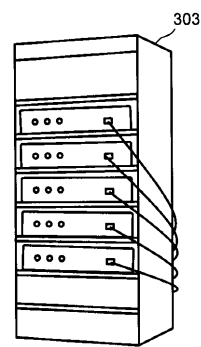


FIG. 33

【図 34】

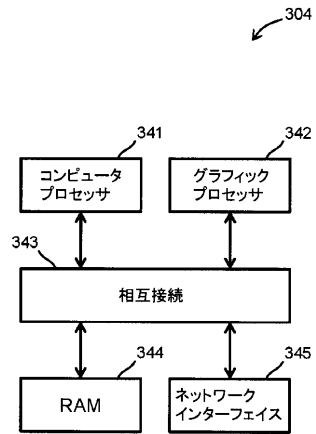


FIG. 34

フロントページの続き

審査官 山下 剛史

- (56)参考文献 河田 岳大 外3名, "両方向N - g r a m確率を用いた誤り文字検出法", 電子情報通信学会論文誌 (J 8 8 - D - I I), 日本, 社団法人電子情報通信学会, 2 0 0 5 年 3 月 1 日, 第J88-D-II巻第3号, p.629-635
荒牧 英治 外2名, "TYPO Writer ヒトはどのように打ち間違えるのか?", 言語処理学会第16回年次大会発表論文集, 日本, 言語処理学会, 2 0 1 0 年 3 月 8 日, p.966-969

(58)調査した分野(Int.Cl., D B名)

G 1 0 L 1 5 / 0 0 - 1 5 / 3 4