

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6789303号
(P6789303)

(45) 発行日 令和2年11月25日(2020.11.25)

(24) 登録日 令和2年11月5日(2020.11.5)

(51) Int.Cl.

F 1

G 06 F 40/216 (2020.01)

G 06 F 40/216

G 06 N 3/04 (2006.01)

G 06 N 3/04

154

G 10 L 15/16 (2006.01)

G 10 L 15/16

請求項の数 14 (全 16 頁)

(21) 出願番号	特願2018-539314 (P2018-539314)	(73) 特許権者	502208397 グーグル エルエルシー
(86) (22) 出願日	平成29年3月17日 (2017.3.17)		アメリカ合衆国 カリフォルニア州 94 043 マウンテン ビュー アンフィシ
(65) 公表番号	特表2019-511033 (P2019-511033A)		アター パークウェイ 1600
(43) 公表日	平成31年4月18日 (2019.4.18)	(74) 代理人	100108453 弁理士 村山 靖彦
(86) 國際出願番号	PCT/US2017/023047	(74) 代理人	100110364 弁理士 実広 信哉
(87) 國際公開番号	W02017/161320	(74) 代理人	100133400 弁理士 阿部 達彦
(87) 國際公開日	平成29年9月21日 (2017.9.21)	(72) 発明者	ユアン・チャン アメリカ合衆国・カリフォルニア・940 43・マウンテン・ビュー・アンフィシア ター・パークウェイ・1600
審査請求日	平成30年8月13日 (2018.8.13)		最終頁に続く
(31) 優先権主張番号	62/310,513		
(32) 優先日	平成28年3月18日 (2016.3.18)		
(33) 優先権主張国・地域又は機関	米国(US)		

前置審査

(54) 【発明の名称】ニューラルネットワークを用いたテキストセグメントの係り受け解析の生成

(57) 【特許請求の範囲】

【請求項1】

1つまたは複数のコンピュータによって実装されるシステムであって、
入力テキストシーケンス内の各トークンについて、
前記トークンについての特徴入力を受け取り、
前記トークンについての前記特徴入力を集合的に処理して、前記入力テキストシーケンス内の前記トークンの品詞を決定する際に使用するための、前記特徴入力の代替表現を生成する

ように構成された1つまたは複数の初期ニューラルネットワークレイヤ
を備える第1のニューラルネットワークと、

前記1つまたは複数の初期ニューラルネットワークレイヤによって生成された前記入力
テキストシーケンス内の前記トークンについての前記特徴入力の前記代替表現を処理して
、前記代替表現を使用してパーサ状態を反復的に更新することによって前記入力テキスト
シーケンスの係り受け解析を生成することであって、前記係り受け解析を生成する際に前
記入力テキストシーケンス内の前記トークンの前記品詞が直接使用されない、ことを
行うように構成された係り受け解析ニューラルネットワークと
を備えるシステム。

【請求項2】

前記第1のニューラルネットワークが、
各トークンについて、

10

20

前記特徴入力の前記代替表現を受け取り、対応する品詞が前記入力テキストシーケンス内の前記トークンの前記品詞である可能性を表す、複数の品詞タグのそれぞれについてのそれぞれのスコアを生成する

ように構成された出力レイヤ

をさらに備える請求項1に記載のシステム。

【請求項3】

前記係り受け解析ニューラルネットワークが状態ベースの係り受け解析ニューラルネットワークである請求項1または2に記載のシステム。

【請求項4】

前記係り受け解析ニューラルネットワークが、現在のパーサ状態での特定の位置のトークンについての前記特徴入力の前記代替表現を処理することによって前記パーサ状態を更新することによって前記係り受け解析を生成するように構成される請求項3に記載のシステム。

10

【請求項5】

前記係り受け解析ニューラルネットワークが、現在のパーサ状態での特定の位置のトークンについての前記特徴入力の前記代替表現を含む状態特徴入力を受け取り、前記状態特徴入力を処理して、複数のパーサ状態更新のそれぞれについてのそれぞれのスコアを生成するように構成される請求項4に記載のシステム。

【請求項6】

前記係り受け解析ニューラルネットワークが、前記係り受け解析を生成する際に前記入力テキストシーケンス内の前記トークンの前記品詞を直接的に使用しない請求項1から5のいずれか一項に記載のシステム。

20

【請求項7】

前記生成した係り受け解析を入力として使用して前記入力テキストシーケンスを解析し、

前記解析した入力テキストシーケンスに基づいて、前記入力テキストシーケンスに応答する動作の実施を引き起こす

ように構成された自然言語処理システムを備える請求項1から6のいずれか一項に記載のシステム。

【請求項8】

30

前記1つまたは複数のコンピュータによって実行されるとき、請求項1から7のいずれか一項に記載のニューラルネットワークシステムを前記1つまたは複数のコンピュータに実装させる命令と共に符号化された1つまたは複数のコンピュータ可読記憶媒体。

【請求項9】

請求項2から7のいずれか一項に記載のシステムの前記第1のニューラルネットワークおよび前記係り受け解析ニューラルネットワークをトレーニングするための方法であって、

トレーニングテキストシーケンスと、各トレーニングテキストシーケンスについて、前記トレーニングテキストシーケンス内の各トークンについての品詞タグのそれぞれとを含む第1のトレーニングデータを得るステップと、

トレーニングテキストシーケンスと、各トレーニングテキストシーケンスについて、前記トレーニングテキストシーケンスの係り受け解析を定義するパーサ状態のセットとを含む第2のトレーニングデータを得るステップと、

40

前記第1および第2のトレーニングデータに関して前記第1のニューラルネットワークおよび前記係り受け解析ニューラルネットワークをトレーニングするステップを含み、トレーニングする前記ステップが、

前記第1のトレーニングデータからのトレーニング例上の前記第1のニューラルネットワークをトレーニングして、前記第1のニューラルネットワークのパラメータの値を更新するステップと、

前記係り受け解析ニューラルネットワークおよび前記第1のニューラルネットワークの前記初期ニューラルネットワークレイヤをトレーニングするステップであって、

50

前記第2のトレーニングデータからのトレーニング例について解析損失を決定することと、

前記係り受け解析ニューラルネットワークおよび前記初期ニューラルネットワークレイヤを通じて前記解析損失を逆伝播し、前記係り受け解析ニューラルネットワークおよび前記初期ニューラルネットワークレイヤのパラメータの値を更新することと
を含むステップと

を反復的に実施するステップを含む、方法。

【請求項 1 0】

反復的に実施する前記ステップが、

前記第1のニューラルネットワークの前記トレーニングを実施するステップと、前記第1 10
および第2のトレーニングデータとは異なるトレーニング例に関して前記解析ニューラル
ネットワークおよび前記初期ニューラルネットワークレイヤの前記トレーニングとを実施
するステップとを反復的に交替するステップ
を含む請求項9に記載の方法。

【請求項 1 1】

前記第2のトレーニングデータを得る前記ステップが、

前記第2のトレーニングデータ内の前記トレーニングテキストシーケンスのそれぞれに
ついて、

前記トレーニングテキストシーケンスの前記係り受け解析を得るステップと、

前記係り受け解析を広げて、前記トレーニングテキストシーケンスの前記係り受け解
析を定義するパーサ状態の前記セットを決定するステップと
を含む請求項9または10に記載の方法。

【請求項 1 2】

前記第1および第2のトレーニングデータに関して前記第1のニューラルネットワークお
よび前記係り受け解析ニューラルネットワークをトレーニングする前記ステップが、

前記トレーニング動作を反復的に実施する前に、前記第1のニューラルネットワークを
事前トレーニングするステップ

をさらに含む請求項9から11のいずれか一項に記載の方法。

【請求項 1 3】

1つまたは複数のコンピュータによって実行されるとき、請求項9から12のいずれか一項
に記載の方法の動作を前記1つまたは複数のコンピュータに実施させる命令と共に符号化
された1つまたは複数のコンピュータ可読記憶媒体。

【請求項 1 4】

1つまたは複数のコンピュータと、前記1つまたは複数のコンピュータによって実行され
るとき、請求項9から12のいずれか一項に記載の方法の動作を前記1つまたは複数のコンピ
ュータに実施させる命令を記憶する1つまたは複数の記憶デバイスとを備えるシステム。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本明細書は、ニューラルネットワークを使用して、テキストセグメントについての係り
受け解析(dependency parse)を生成することに関する。

【背景技術】

【0 0 0 2】

ニューラルネットワークは、非線型ユニットの1つまたは複数のレイヤを利用して、受け取った入力についての出力を予測する機械学習モデルである。いくつかのニューラルネットワークは、出力レイヤに加えて1つまたは複数の隠れレイヤを含む。各隠れレイヤの出力は、ネットワーク内の次のレイヤ、すなわち次の隠れレイヤまたは出力レイヤに対する入力として使用される。ネットワークの各レイヤは、パラメータのそれぞれのセットの現在値に従って、受け取った入力から出力を生成する。

【先行技術文献】

10

20

30

40

50

【非特許文献】**【0003】**

【非特許文献1】http://nlp.stanford.edu/software/dependencies_manual.pdf

【非特許文献2】Danqi ChenおよびChristopher Manning、A Fast and Accurate Dependency Parser Using Neural Networks. In Proceedings of EMNLP 2014

【発明の概要】**【発明が解決しようとする課題】****【0004】**

自然言語処理は、コンピュータと人間の(自然)言語との間の対話に関する計算機科学の一分野である。自然言語処理は、人間機械自然言語インターフェースを提供する際に使用されることを含む多くの適用分野を有する。そのようなインターフェースは、人間が、例えば、自分の音声のみを使用して「スマート環境」においてデバイスを制御することを可能にし得る。

【0005】

音声コマンドまたは発声だけに基づいて人間の意図を理解することは、データ集約的なタスクであり得、正確に実施することが困難であり得る。したがって、音声コマンドの意図を理解することは、多くの場合不可能であり、したがって、コマンドを理解しようと試みる間に計算資源を使い果たしたにも関わらず、所期の動作(例えば、特定のデバイスの、特定の方式での制御)が正しく実施されず、または全く実施されないことがある。

【0006】

したがって、そのような状況の発生を低減するために、自然言語処理の信頼性を改善することが望ましい。自然言語処理の速度を向上させ、かつ/または計算資源の使用量を削減することも望ましいことがある。

【課題を解決するための手段】**【0007】**

自然言語処理において使用され得る1つの入力は係り受け解析であり、一般的に言えば、係り受け解析は、テキストセグメント内のワード間の依存関係を定義する。すなわち、係り受け解析は、テキストセグメント内のどのワードがテキストセグメント内の他のどのワードに依存するか、および任意選択で、依存関係のそれぞれのタイプを定義する。本明細書は、1つまたは複数の位置における1つまたは複数のコンピュータ上のコンピュータプログラムとして実装されたシステムが、改良型の方式で入力テキストセグメントについての係り受け解析をどのように生成することができ、それによって、入力テキストセグメントに関する自然言語処理の性能を潜在的に改善するかを説明する。

【0008】

本明細書において説明される主題の特定の実施形態は、以下の利点のうちの1つまたは複数を実現するように実装され得る。テキストセグメントの係り受け解析を生成する際にPOSタグを直接的に使用しないことにより、テキストセグメントの係り受け解析を予測する際のPOSタグ付け誤りの連鎖が低減され、係り受け解析が正確に予測され得る。語彙化された特徴を係り受けパーサニューラルネットワークによって使用することを必要としないことにより、その代わりに初期ニューラルネットワークレイヤによって生成された表現を再利用することにより、係り受け解析を生成するモデルはより小さいサイズを有し、すなわちよりコンパクトであり、語彙化された特徴を使用する代替よりも高速に係り受け解析を生成し得る。理解するであろうが、より迅速かつ/または正確に係り受け解析を生成することにより、入力テキストセグメントおよび係り受け解析に基づいて自然言語処理システムによって決定される応答動作(例えば、スマートホームでのデバイスの制御)も、より迅速かつ/または正確に実施され得る。さらに、複数のタスクにわたる、すなわちPOSタグ付けおよび係り受け解析にわたるパイプラインまでの逆伝播を可能にする連続的な形のスタッキングを使用することにより、モデルの構成要素が効率的かつ迅速にトレーニングされ得る。

【0009】

10

20

30

40

50

本明細書において説明される主題の1つまたは複数の実施形態の詳細が、添付の図面および以下の説明において述べられる。説明、図面、および特許請求の範囲から主題の他の特徴、態様、および利点が明らかとなるであろう。

【図面の簡単な説明】

【0010】

【図1】例示的な係り受け解析生成システムを示す図である。

【図2】入力テキストセグメントについての係り受け解析を生成するための例示的なプロセスの流れ図である。

【図3】現在の解析状態を更新するための例示的なプロセスの流れ図である。

【図4】POSタグ付けニューラルネットワークおよび係り受け解析ニューラルネットワークをトレーニングするための例示的なプロセスの流れ図である。 10

【発明を実施するための形態】

【0011】

様々な画面内の同様の参照番号および名称は、同様の要素を示す。

【0012】

図1は例示的な係り受け解析生成システム100を示す。係り受け解析生成システム100は、1つまたは複数の位置における1つまたは複数のコンピュータ上のコンピュータプログラムとして実装されたシステムの一例であり、その中で、以下で説明されるシステム、構成要素、および技法が実装され得る。

【0013】

係り受け解析生成システム100は、入力テキストセグメント102を受け取り、各入力テキストセグメントについてのそれぞれの係り受け解析152を生成する。各入力テキストセグメントは、トークンのシーケンス、すなわち特定の自然言語のワードおよび任意選択で句読点のシーケンスである。例えば、セグメントは、センテンス、センテンスフラグメント、または別の複数ワードシーケンスであり得る。 20

【0014】

所与のテキストセグメントについての係り受け解析は、セグメント内のトークン間の依存関係を定義するデータである。具体的には、係り受け解析は、セグメント内のヘッドワードと、ヘッドワードに依存するセグメント内のワードおよび任意選択で句読点とを特定する。 30

【0015】

任意選択で、係り受け解析はまた、所与のヘッドワードに依存する各ワードについて、ヘッドワードと従属ワードとの間の関係のタイプ、例えば従属ワードはヘッドワードの形容詞補語か、ヘッドワードの副詞修飾語か、ヘッドワードの直接目的語か、などを指定するラベルをも含み得る。係り受け解析においてトークン間で指定され得る関係タイプの例示的なセットが、http://nlp.stanford.edu/software/dependencies_manual.pdfにおいて説明されている。

【0016】

いくつかのケースでは、係り受け解析はトリプルのセットであり、各トリプルは、セグメントからヘッドワードを特定し、セグメントから従属ワードを特定し、ヘッドワードと従属ワードとの間の関係のタイプを指定するラベルを特定する。 40

【0017】

係り受け解析生成システム100は、1つまたは複数の初期ニューラルネットワークレイヤ110、解析サブシステム130、および係り受け解析ニューラルネットワーク140を含む。トレーニングの間、および任意選択でトレーニングの後、係り受け解析生成システム100はまた、POS(part-of-speech)出力レイヤ120をも含む。

【0018】

入力テキストセグメント102についての係り受け解析を生成する部分として、係り受け解析生成システム100は、初期ニューラルネットワークレイヤ110を使用して、入力テキストセグメント102内の各トークンについてのそれぞれの特徴入力を処理する。 50

【 0 0 1 9 】

一般には、所与のトークンについての特徴入力は、その所与のトークンの特徴と、入力テキストセグメント内のその所与のトークンを取り囲む1つまたは複数のトークン、すなわち入力テキストセグメント内のその所与のトークンのウィンドウ内のトークンの特徴とを含む。

【 0 0 2 0 】

トークンの特徴は、トークンを特徴付けるベクトルまたは他の数値の集合である。例えば、トークンの特徴は、トークンを特定する、すなわちワードまたは句読点を特定するワンホット符号化特徴ベクトル(one-hot encoded feature vector)、トークンがシンボル、例えばハイフン、数字、または句読点を含むかどうかを示す特徴ベクトル、トークンが接頭辞または接尾辞を含むワードであるかどうかを示す特徴ベクトル、トークンが大文字化されたワードであるかどうかを示す特徴ベクトルなどを含み得る。

10

【 0 0 2 1 】

いくつかの実装では、ウィンドウは各特徴について同一の固定サイズである。他の実装では、異なる特徴は異なるウィンドウサイズを有する。例えば、トークンを特定するワンホット符号化特徴ベクトルについてのウィンドウサイズは、所与のトークンの3つのトークン以内のトークンを含み得、大文字化を特定する特徴ベクトルは、所与のトークンの1つのトークン以内のトークンを含み得る。

【 0 0 2 2 】

初期ニューラルネットワークレイヤ110は、入力セグメント内の各トークンについて、トークンについての特徴入力を処理して、トークンについての特徴表現112を生成するよう²⁰に集合的に構成されるフィードフォワードニューラルネットワークレイヤである。トークンについての特徴表現は、トークンについての特徴入力の代替表現、すなわちトークンについての特徴入力を表すベクトルまたは他の数値の集合である。

【 0 0 2 3 】

例えば、初期ニューラルネットワークレイヤ110は、埋込みレイヤと、その後に続く1つまたは複数の完全に接続された隠れレイヤとを含み得る。

【 0 0 2 4 】

埋込みレイヤは、一般にはまばらである、特徴入力内の特徴を受け取り、各特徴についてのそれぞれの埋込みベクトルを決定し、埋込みベクトルを組み合わせ、例えば連結し、特徴の組合せ埋込みを生成するように構成される。

30

【 0 0 2 5 】

1つまたは複数の完全に接続された隠れレイヤは、組合せ埋込みを処理して、すなわち1つまたは複数の非線形変換を組合せ埋込みに適用することによって処理して、トークンについての特徴表現を生成するように構成される。

【 0 0 2 6 】

解析サブシステム130は、初期ニューラルネットワークレイヤ110によって生成された入力テキストセグメント102内のトークンについての特徴表現112を受け取り、特徴表現112を使用して、係り受け解析ニューラルネットワーク140についての入力を生成する。

40

【 0 0 2 7 】

係り受け解析ニューラルネットワーク140は、状態特徴132、すなわち解析サブシステム130によって維持される解析状態の現在の構成の特徴を受け取り、現在の解析状態に対する状態更新142を定義する出力を生成するように構成されるフィードフォワードニューラルネットワークである。いくつかの実装では、係り受け解析ニューラルネットワーク140は、埋込みレイヤ、1つまたは複数の完全に接続された隠れレイヤ、および出力レイヤを含む。

【 0 0 2 8 】

係り受け解析の生成中の任意の所与の時刻で、解析サブシステム130によって維持される解析状態は、スタッカまたはバッファ内のそれぞれの位置を、入力シーケンス内のトークンの一部またはすべてに割り当て、入力セグメントについての、それまでに生成された

50

依存関係トリプルを特定する。解析状態の例示的な構成が、Danqi ChenおよびChristophe Manning、A Fast and Accurate Dependency Parser Using Neural Networks. In Proceedings of EMNLP 2014においてより詳細に説明されている。

【0029】

所与の構成の状態特徴は、複数の特徴表現の順序付き組合せ、例えば行列または連結ベクトルであり、順序付き組合せ内の各位置は、解析状態のスタックまたはバッファ内の異なる位置に対応する。

【0030】

いくつかの実装では、解析状態の現在の構成についての状態特徴を生成するために、解析サブシステム130は、順序付き組合せ内の各位置での特徴表現として、解析状態の現在の構成内の対応する位置にあるトークンの特徴表現を選択する。したがって、順序付き組合せ内の各位置は、解析状態の現在の構成内の対応する位置からの特徴表現を含む。簡略化した例として、順序付き組合せ内の位置が[buffer_0;stack_0;stack_1]であり、現在の構成がバッファ内の位置0においてtoken1を含み、スタック内の位置0においてtoken2を含み、スタック内の位置1においてtoken3を含む場合、現在の構成についての状態特徴は、[token1についての特徴表現, token2についての特徴表現, token3についての特徴表現]となる。順序付き組合せ内の位置のうちの1つに対応する位置が空である場合、解析サブシステム130は、指定のヌル特徴表現を順序付き組合せに追加し得る。

10

【0031】

いくつかの実装では、順序付き組合せ内の各位置は、解析状態の現在の構成内の対応する位置を有し、順序付き組合せ位置での特徴は、対応する構成位置にあるトークンの特徴表現と、それまでに生成された依存関係トリプルによる対応する構成位置にあるトークンに関するトークン、例えば子トークンおよび兄弟トークンの特徴表現の組合せ、例えば連結である。対応する構成位置のトークンが、それまでに生成されたトリプル内の特定の関係を有するトークンを有さない場合、システムは指定のヌル特徴表現を使用し得る。

20

【0032】

いくつかの実装では、順序付き組合せはまた、それまでに生成された依存関係トリプルについてのラベルの数値表現をも含む。

【0033】

解析サブシステム130および係り受け解析ニューラルネットワーク140は、解析状態を反復的に更新することによって、特徴表現112から係り受け解析152を生成する。係り受け解析を生成すること、および解析状態を更新することは、図2および図3を参照しながら以下により詳細に説明される。

30

【0034】

POS出力レイヤ120は、入力テキストセグメント内の所与のトークンの特徴表現112を処理して、その所与のトークンについてのpart-of-speech割当てを定義する、その所与のトークンについてのPOSタグ122を定義する出力を生成するように構成される。具体的には、POS出力レイヤ120は、特徴表現112を処理して、POSタグの所定のセット内の各POSタグについてのそれぞれのスコアを生成するように構成されるsoftmaxまたは他の種類のニューラルネットワーク出力レイヤである。各スコアは、対応する品詞が入力テキストシーケンス内のトークンの品詞である可能性を表す。

40

【0035】

したがって、特徴表現112は、入力セグメント102内のトークンについてのPOSタグを生成し、すなわち最高のスコアのPOSタグをトークンについてのPOSタグとして選択することによって生成し、かつ入力セグメント102の係り受け解析152を生成するように使用され得る。しかしながら、係り受け解析ニューラルネットワーク140は、係り受け解析152を生成するとき、トークンについてのPOSタグを直接的に使用せず、すなわち特徴表現112のみを使用することによって使用しない。

【0036】

いくつかの実装では、生成された後に、係り受け解析生成システム100は、さらなる処

50

理のために、係り受け解析152を別のシステムに対する入力として供給する。例えば、係り受け解析生成システム100は、入力テキストセグメント102を解析し、応答動作の実施を引き起こす自然言語処理システムに対する入力として係り受け解析152を供給し得る。例えば、自然言語処理システムは、人間機械自然言語(例えば音声)インターフェースを容易にするために利用され得る。そのような例では、応答動作は、スマート環境内のデバイスを制御することを含み得る。自然言語処理システムによって実施され得る動作の他の例には、感情解析、質問回答、要約、または他の自然言語処理タスクが含まれる。例えば、これらの動作は、例えばスマートスピーカまたは他の自動インテリジェントパーソナルアシスタントサービスによって要求される、処理ユーザ要求の部分として実施され得る。

【0037】

10

別のシステムに対する入力として係り受け解析152を供給する代わりに、またはそれに加えて、係り受け解析生成システム100は、係り受け解析生成システム100のユーザへの提示のために、係り受け解析を記憶し、または係り受け解析を指定するデータを供給し得る。

【0038】

20

係り受け解析生成システム100は、初期ニューラルネットワークレイヤ110および係り受け解析ニューラルネットワーク140をトレーニングして、初期ニューラルネットワークレイヤ110および係り受け解析ニューラルネットワーク140のパラメータのトレーニング後の値を決定し得る。初期ニューラルネットワークレイヤおよび係り受け解析ニューラルネットワークをトレーニングすることは、図4を参照しながら以下でより詳細に説明される。

【0039】

図2は、入力テキストシーケンスについての係り受け解析を生成するための例示的なプロセス200の流れ図である。便宜上、プロセス200は、1つまたは複数の位置に配置された1つまたは複数のコンピュータのシステムによって実施されるものとして説明される。例えば、適切にプログラムされた係り受け解析生成システム、例えば図1の係り受け解析生成システム100が、プロセス200を実施し得る。

【0040】

システムは入力テキストセグメントを得る(ステップ202)。入力テキストセグメントは、入力順に従って配置された複数のトークン、すなわちワードおよび任意選択で句読点を含む。

30

【0041】

システムは、1つまたは複数の初期ニューラルネットワークレイヤを通じて、入力テキストセグメント内のトークンのそれぞれについての特徴入力を処理する(ステップ204)。前述のように、初期ニューラルネットワークレイヤは、所与のトークンの特徴入力を処理して、トークンの品詞を決定する際に使用するために、すなわちPOSタグ出力レイヤによって処理するために、および入力テキストセグメントの係り受け解析を生成する際に使用するために、特徴表現、すなわち特徴の代替表現を生成するように構成される。

【0042】

システムは、係り受け解析ニューラルネットワークを使用して代替表現を処理して、入力テキストセグメントの係り受け解析を生成する(ステップ206)。

40

【0043】

具体的には、システムは解析状態を初期化し、係り受け解析ニューラルネットワークを使用して解析状態を反復的に更新し、係り受け解析を生成する。初期解析状態は、例えば、スタック内の所定のROOTトークンと、入力セグメント内のその位置に従って配置されたバッファ内の入力セグメント内のトークンとを含み得、既に生成された依存関係トリプルを含まない。システムは、解析状態が終端状態となるまで、解析状態を反復的に更新し、次いで終端状態の依存関係トリプルを、入力テキストセグメントについての係り受け解析として扱う。例えば、終端状態は、空バッファと、スタック内のROOTトークンのみを含み得る。

【0044】

50

解析状態を更新することは、図3を参照しながら以下でより詳細に説明される。

【0045】

図3は、解析状態を更新するための例示的なプロセス300の流れ図である。便宜上、プロセス300は、1つまたは複数の位置に配置された1つまたは複数のコンピュータのシステムによって実施されるものとして説明される。例えば、適切にプログラムされた係り受け解析生成システム、例えば図1の係り受け解析生成システム100が、プロセス300を実施し得る。

【0046】

システムは、現在の解析状態を使用して状態特徴を生成する(ステップ302)。具体的には、システムは、解析状態の現在の構成内の対応する位置にあるトークンについての特徴表現を、順序付き組合せ内の各位置に追加することによって、特徴表現の順序付き組合せを生成する。10

【0047】

システムは、係り受け解析ニューラルネットワークを使用して状態特徴を処理する(ステップ304)。係り受け解析ニューラルネットワークは、状態特徴を処理して、現在の解析状態に対する更新を定義する出力を生成するように構成される。

【0048】

より具体的には、係り受け解析ニューラルネットワークは、潜在的解析状態更新のセット内の各潜在的更新についてのそれぞれのスコアを生成する。解析状態更新は、バッファからトークンを除去し、トークンをスタック上にプッシュする更新を含み、従属ワードとヘッドワードとの間の可能な各ラベルについて、(i)スタック上の第2のトークンを第1のトークンの従属とマークするそれぞれの左アーク更新が、対応するラベルを依存関係に割り当て、スタックから第2のトークンを除去し、(ii)スタック上の第1のトークンを第2のトークンの従属とマークするそれぞれの右アーク更新が、対応するラベルを依存関係に割り当て、スタックから第1のトークンを除去する。例えば、可能なラベルは、http://nlp.stanford.edu/software/dependencies_manual.pdfにおいて説明されている関係タイプのセット内の関係タイプの一部またはすべてでよい。20

【0049】

システムは、出力を使用して現在の解析状態を更新する(ステップ306)。具体的には、システムは、スコアを使用して、すなわち最高のスコアを有する更新を選択することによって、または対応するスコアに従って、可能な更新からサンプリングすることによって、可能な更新から更新を選択し、更新を現在の解析状態に適用して、更新後解析状態を生成する。30

【0050】

更新後解析状態が終端解析状態である場合、システムは、更新後解析状態によって定義される係り受け解析を、テキストセグメントについての最終係り受け解析として選択する。。

【0051】

更新後解析状態が所定の終端解析状態ではない場合、システムは、現在の解析状態の代わりに更新後解析状態でプロセス300を反復する。40

【0052】

図4は、POSタグ付けニューラルネットワークおよび係り受け解析ニューラルネットワークをトレーニングするための例示的なプロセス400の流れ図である。便宜上、プロセス400は、1つまたは複数の位置に配置された1つまたは複数のコンピュータのシステムによって実施されるものとして説明される。例えば、適切にプログラムされた係り受け解析生成システム、例えば図1の係り受け解析生成システム100が、プロセス400を実施し得る。

【0053】

システムは、プロセス400を実施して、POSタグ付けニューラルネットワーク、すなわち図1の1つまたは複数の初期ニューラルネットワークレイヤおよびPOS出力レイヤを含むニューラルネットワークと、係り受け解析ニューラルネットワーク、すなわち図1の係り受50

け解析ニューラルネットワークとをトレーニングする。

【0054】

システムは、トレーニングテキストシーケンスと、各トレーニングテキストシーケンスについて、トレーニングテキストシーケンス内の各トークンについてのそれぞれのPOSタグとを含む第1のトレーニングデータを得る(ステップ402)。

【0055】

システムは、トレーニングテキストシーケンスと、各トレーニングテキストシーケンスについて、トレーニングテキストシーケンスの係り受け解析を定義するパーサ状態のセットとを含む第2のトレーニングデータを得る(ステップ404)。

【0056】

すなわち、第2のトレーニングデータ内の所与のトレーニングテキストシーケンスについて、システムは、係り受け解析を生成する(パーサ状態、アクション)対のセットを得る。各(パーサ状態、アクション)対内のアクションは、トレーニングテキストシーケンスについて正しい係り受け解析を生成するために対の中のパーサ状態に適用すべき状態更新である。

【0057】

いくつかの実装では、システムは(パーサ状態、アクション)対を受け取る。他の実装では、システムは、シーケンスについての係り受け解析を受け取り、係り受け解析を広げて、シーケンスについての(パーサ状態、アクション)対を決定する。

【0058】

次いで、システムは、第1および第2のトレーニングデータに関してPOSニューラルネットワークおよび係り受け解析ニューラルネットワークをトレーニングして、初期ニューラルネットワークレイヤ、POS出力レイヤ(出力レイヤがトレーニング可能なパラメータを有する場合)、および係り受け解析ニューラルネットワークのパラメータのトレーニング後の値を決定する(ステップ306)。

【0059】

具体的には、システムは、以下を反復的に実施することによってネットワークをトレーニングする。(i)第1のトレーニングデータからのトレーニング例に関してPOSニューラルネットワークをトレーニングして、POSニューラルネットワークのパラメータの値を更新すること、および(ii)第2のトレーニングデータからのトレーニング例に関して解析ニューラルネットワークおよび初期ニューラルネットワークレイヤをトレーニングすること。

【0060】

第1のトレーニングデータからのトレーニング例は、第1のトレーニングデータ内のトレーニングシーケンスのうちの1つからのトークンについての特徴入力、およびトークンについての対応するPOSタグである。

【0061】

第2のトレーニングデータからのトレーニング例は、第2のトレーニングデータ内の(解析状態、アクション)対に対応する状態特徴、および(解析状態、アクション)対からのアクションである。

【0062】

システムは、動作(i)および(ii)の各反復について、対応するトレーニングデータ内のトレーニング例からランダムにトレーニング例を選択し得る。

【0063】

動作(i)および(ii)を反復的に実施するために、システムは、動作(i)を実施することと、動作(ii)を実施することとの間で反復的に交替し得る。例えば、システムは、第1の数のトレーニングエポックについて動作(i)を実施することと、第2の数のトレーニングエポックについて動作(ii)を実施することとの間で交替し得る。いくつかの実装では、第1の数は第2の数に等しい。いくつかの他の実装では、第1の数は第2の数とは異なる。システムは、トレーニングデータ内のトレーニング例のすべてがサンプリングされるまで、または何らかの他の終了基準が満たされるまで、所定の反復数について動作を実施することの

10

20

30

40

50

間で交替し得る。

【 0 0 6 4 】

第2のトレーニングデータからのトレーニング例に関して解析ニューラルネットワークおよび初期ニューラルネットワークレイヤをトレーニングするために、システムは、トレーニング例についての解析損失の勾配を決定し、係り受け解析ニューラルネットワークおよび初期ニューラルネットワークレイヤを通じて勾配を逆伝播し、係り受け解析ニューラルネットワークおよび初期ニューラルネットワークレイヤのパラメータの値を、すなわち逆伝播トレーニング技法を用いる確率的勾配降下を使用して更新する。

【 0 0 6 5 】

具体的には、システムは、(i)トレーニング例内の状態特徴を処理することによって解析ニューラルネットワークによって生成されたスコアと、(ii)トレーニング例内のアクションを状態に適用すべき状態更新として特定するスコアのセットとの間の誤差から解析損失を決定する。

【 0 0 6 6 】

第1のトレーニングデータからのトレーニング例に関してPOSニューラルネットワークをトレーニングするために、システムは、POS損失の勾配を決定し、POS出力レイヤおよび初期ニューラルネットワークレイヤを通じて勾配を逆伝播し、初期ニューラルネットワークレイヤのパラメータの値を、すなわち逆伝播トレーニング技法を用いる同一の確率的勾配降下、または逆伝播トレーニング技法を用いる異なる確率的勾配降下を使用して更新する。

【 0 0 6 7 】

具体的には、システムは、(i)トレーニング例内のトークンについての特徴入力を処理することによってPOSニューラルネットワークによって生成されたスコアと、(ii)トレーニング例内のPOSタグがトークンについての正しいPOSタグであることを示すスコアのセットとの間の誤差からPOS損失を決定する。

【 0 0 6 8 】

いくつかの実装では、システムは、トレーニング動作を反復的に実施する前に、POSニューラルネットワークを事前トレーニングする。すなわち、システムは、動作(ii)を実施することを開始する前に、動作(i)を反復的に実施することによってPOSニューラルネットワークを事前トレーニングする。

【 0 0 6 9 】

本明細書において説明される主題および機能的動作の実施形態は、本明細書において開示された構造およびその構造的均等物を含む、デジタル電子回路、有形に実施されたコンピュータソフトウェアまたはファームウェア、コンピュータハードウェア、またはそれらのうちの1つまたは複数の組合せとして実装され得る。

【 0 0 7 0 】

本明細書において説明される主題の実施形態は、1つまたは複数のコンピュータプログラム、すなわちデータ処理装置による実行のために、またはデータ処理装置の動作を制御するために有形の非一時的プログラムキャリア上に符号化されたコンピュータプログラム命令の1つまたは複数のモジュールとして実装され得る。代替または追加として、プログラム命令は、データ処理装置による実行のために、適切な受信機装置への送信のために情報を符号化するように生成される、人工的に生成された伝播信号、例えば機械で生成された電気信号、光信号、または電磁信号上で符号化され得る。コンピュータ記憶媒体は、機械可読記憶デバイス、機械可読記憶基板、ランダムもしくはシリアルアクセスメモリデバイス、またはそれらのうちの1つまたは複数の組合せであり得る。

【 0 0 7 1 】

「データ処理装置」という用語は、例としてプログラム可能プロセッサ、コンピュータ、または複数のプロセッサもしくはコンピュータを含む、データを処理するためのすべての種類の装置、デバイス、および機械を包含する。装置は、専用論理回路、例えばFPGA(フィールドプログラマブルゲートアレイ)またはASIC(特定用途向け集積回路)を含み得る

10

20

30

40

50

。装置はまた、ハードウェアに加えて、当該のコンピュータプログラムのための実行環境を生成するコード、例えばプロセッサファームウェア、プロトコルスタック、データベース管理システム、オペレーティングシステム、またはそれらのうちの1つまたは複数の組合せを構成するコードをも含み得る。

【0072】

コンピュータプログラム(プログラム、ソフトウェア、ソフトウェアアプリケーション、モジュール、ソフトウェアモジュール、スクリプト、またはコードとも呼ばれ、またはそれらとして説明されることがある)が、コンパイル型言語またはインタプリタ型言語、宣言型言語または手続型言語を含む任意の形態のプログラミング言語として書かれ得、スタンダロンプログラムとして、またはモジュール、コンポーネント、サブルーチン、もしくはコンピューティング環境内での使用に適した他のユニットとしての形態を含む任意の形態として配置され得る。コンピュータプログラムは、必須ではないがファイルシステム内のファイルに対応し得る。プログラムは、他のプログラムまたはデータを保持するファイルの一部、例えばマークアップ言語文書内に記憶された1つまたは複数のスクリプトの中に、当該のプログラム専用の単一のファイル内に、または複数の協調ファイル、例えば1つまたは複数のモジュール、サブプログラム、もしくはコードの部分を記憶するファイル内に格納され得る。コンピュータプログラムは、1つのコンピュータ上で、または1つの場所に配置され、または複数の場所にわたって分散され、通信ネットワークによって相互接続される複数のコンピュータ上で実行されるように配置され得る。

【0073】

本明細書において説明されたプロセスおよび論理フローは、1つまたは複数のコンピュータプログラムを実行して、入力データに対して演算し、出力を生成することによって機能を実施する1つまたは複数のプログラム可能コンピュータによって実施され得る。プロセスおよび論理フローはまた、専用論理回路、例えばFPGA(フィールドプログラマブルゲートアレイ)またはASIC(特定用途向け集積回路)によっても実施され得、装置は、それらとしても実装され得る。

【0074】

コンピュータプログラムの実行に適したコンピュータは、例として、汎用もしくは専用マイクロプロセッサまたはその両方、あるいは任意の他の種類の中央演算処理装置に基づくものであり得る。一般には、中央演算処理装置は、読み取り専用メモリもしくはランダムアクセスメモリまたはその両方から命令およびデータを受け取る。コンピュータの不可欠な要素は、命令を実施または実行するための中央演算処理装置と、命令およびデータを記憶するための1つまたは複数のメモリデバイスである。一般には、コンピュータはまた、データを記憶するための1つまたは複数の大容量記憶デバイス、例えば磁気ディスク、光磁気ディスク、または光ディスクをも含み、あるいはそれらからデータを受け取り、もしくはそれらにデータを転送し、またはその両方を行うように動作可能に結合される。しかしながら、コンピュータはそのようなデバイスを有する必要はない。さらに、コンピュータは、ほんのいくつかの例を挙げれば、別のデバイス、例えば携帯電話、携帯情報端末(PDA)、モバイルオーディオもしくはビデオプレーヤ、ゲームコンソール、全地球測位システム(GPS)受信機、またはポータブル記憶デバイス、例えばユニバーサルシリアルバス(USB)フラッシュドライブ内に組み込まれ得る。

【0075】

コンピュータプログラム命令およびデータを記憶するのに適したコンピュータ可読媒体は、例として、半導体メモリデバイス、例えばEPROM、EEPROM、およびフラッシュメモリデバイス、磁気ディスク、例えば内部ハードディスクまたは取外し可能ディスク、光磁気ディスク、ならびにCD-ROMおよびDVD-ROMディスクを含む、すべての形態の不揮発性メモリ、媒体、およびメモリデバイスを含む。プロセッサおよびメモリは、専用論理回路によって補足され、またはその中に組み込まれ得る。

【0076】

ユーザとの対話を実現するために、本明細書において説明される主題の実施形態は、ユ

10

20

30

40

50

ーザに情報を表示するためのディスプレイデバイス、例えばCRT(陰極線管)もしくはLCD(液晶ディスプレイ)モニタと、ユーザがそれによってコンピュータに入力を与えることのできるキーボードおよびポインティングデバイス、例えばマウスもしくはトラックボールとを有するコンピュータ上に実装され得る。他の種類のデバイスもユーザとの対話を実現するために使用され得、例えば、ユーザに提供されるフィードバックは、任意の形態の感覚フィードバック、例えば視覚フィードバック、音声フィードバック、または触覚フィードバックであり得、ユーザからの入力は、音響、音声、または触覚入力を含む任意の形態として受け取られ得る。さらに、コンピュータは、ユーザによって使用されるデバイスに文書を送り、デバイスから文書を受け取ることによって、例えば、ウェブブラウザから受け取った要求に応答して、ユーザのクライアントデバイス上のウェブブラウザにウェブページを送ることによって、ユーザと対話し得る。

【0077】

本明細書において説明される主題の実施形態は、バックエンド構成要素を例えればデータサーバとして含み、またはミドルウェア構成要素、例えばアプリケーションサーバを含み、またはフロントエンド構成要素、例えばユーザが本明細書において説明される主題の実装とそれを通じて対話することができるグラフィカルユーザインターフェースもしくはウェブブラウザを有するクライアントコンピュータを含み、あるいは1つまたは複数のそのようなバックエンド、ミドルウェア、またはフロントエンド構成要素の任意の組合せを含むコンピューティングシステムとして実装され得る。システムの構成要素は、任意の形態または媒体のデジタルデータ通信、例えば通信ネットワークによって相互接続され得る。通信ネットワークの例には、ローカルエリアネットワーク(「LAN」)および広域ネットワーク(「WAN」)、例えばインターネットが含まれる。

【0078】

コンピューティングシステムはクライアントおよびサーバを含み得る。クライアントおよびサーバは、一般には互いに離れており、通常は通信ネットワークを通じて対話する。クライアントとサーバの関係は、それぞれのコンピュータ上で実行中の、互いにクライアント-サーバ関係を有するコンピュータプログラムによって生じる。

【0079】

本明細書は、多くの特定の実装詳細を含むが、これらを、何らかの発明の範囲、または特許請求され得るものとの範囲に関する限定と解釈すべきではなく、むしろ特定の発明の特定の実施形態に特有のものであり得る特徴の説明と解釈すべきである。本明細書において別々の実施形態の状況において説明されるいくつかの特徴はまた、単一の実施形態において組み合わせて実装され得る。逆に、単一の実施形態の状況において説明される様々な特徴はまた、複数の実施形態として別々に、または任意の適切な部分組合せとして実装され得る。さらに、特徴が、いくつかの組合せとして働くものとして上記で説明され、さらにはそのように最初に特許請求され得るが、あるケースでは、特許請求される組合せからの1つまたは複数の特徴が組合せから削除され得、特許請求される組合せは、部分組合せまたは部分組合せの変形を対象とし得る。

【0080】

同様に、図面では動作が特定の順序で示されるが、このことは、望ましい結果を達成するため、そのような動作が図示される特定の順序で、または順次的に実施されること、またはすべての図示される動作が実施されることを必要とすると理解すべきではない。いくつかの環境では、マルチタスキングおよび並列処理が有利であることがある。さらに、前述の実施形態での様々なシステムモジュールおよび構成要素の分離を、すべての実施形態においてそのような分離を必要とすると理解すべきではなく、記載のプログラム構成要素およびシステムは一般に、単一のソフトウェア製品として共に一体化され、または複数のソフトウェア製品としてパッケージ化され得ることを理解されたい。

【0081】

主題の特定の実施形態が説明された。他の実施形態は以下の特許請求の範囲内にある。例えば、請求項に記載の動作は、異なる順序で実施され、それでもなお望ましい結果を達

10

20

30

40

50

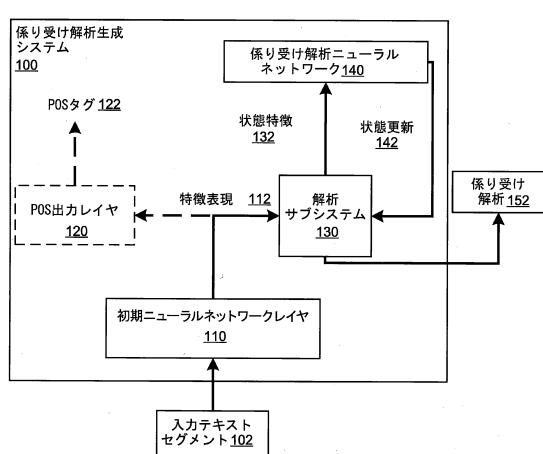
成し得る。一例として、添付の図に示されるプロセスは、望ましい結果を達成するために、図示される特定の順序、または順次的順序を必ずしも必要とするわけではない。いくつかの実装では、マルチタスキングおよび並列処理が有利であることがある。

【符号の説明】

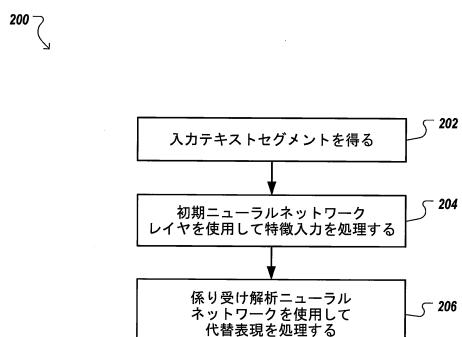
【0082】

204 ノード	
100 係り受け解析生成システム	
102 入力テキストセグメント	
110 初期ニューラルネットワークレイヤ	
112 特徴表現	10
120 POS出力レイヤ	
122 POSタグ	
130 解析サブシステム	
132 状態特徴	
140 係り受け解析ニューラルネットワーク	
142 状態更新	
152 係り受け解析	

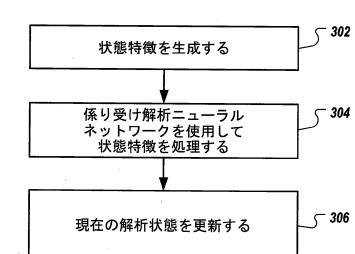
【図1】



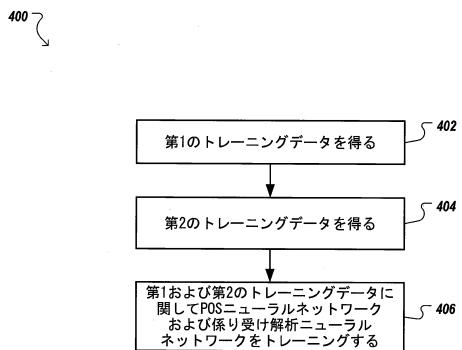
【図2】



【図3】



【図4】



フロントページの続き

(72)発明者 ディヴィッド・ジョセフ・ワイズ

アメリカ合衆国・カリフォルニア・94043・マウンテン・ビュー・アンフィシアター・パーク
ウェイ・1600

審査官 野村 和史

(56)参考文献 特開平03-263266(JP,A)

米国特許出願公開第2011/0301942(US,A1)

中国特許出願公開第1313972(CN,A)

Natural Language Processing (Almost) from Scratch, 2011年, URL, <http://www.jmlr.org/papers/volume12/colllobert11a/colllobert11a.pdf>

Jun Hatori 他3名, Incremental Joint Approach to Word Segmentation, POS Tagging, and Dependency Parsing in Chinese, Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics, 2012年, 1045~1053, URL, <http://wing.comp.nus.edu.sg/~antho/P/P12/P12-1110.pdf>

Danqi Chen et al, A Fast and Accurate Dependency Parser using Neural Networks, 2014年, URL, <http://nlp.stanford.edu/pubs/emnlp2014-depparser.pdf>

(58)調査した分野(Int.Cl., DB名)

G 06 F 40/00 - 40/58

G 06 N 3/04

G 10 L 15/16