

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2007年2月1日 (01.02.2007)

PCT

(10) 国際公開番号
WO 2007/013521 A1

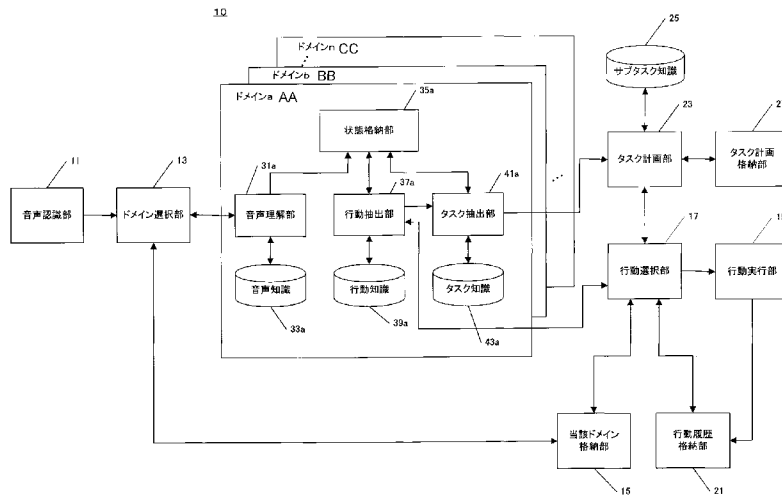
- (51) 国際特許分類:
G10L 15/22 (2006.01) G06F 17/28 (2006.01)
- (21) 国際出願番号: PCT/JP2006/314786
- (22) 国際出願日: 2006年7月26日 (26.07.2006)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
60/703,050 2005年7月26日 (26.07.2005) US
- (71) 出願人 (米国を除く全ての指定国について): 本田技研工業株式会社 (HONDA MOTOR CO., LTD.) [JP/JP]; 〒1078556 東京都港区南青山二丁目1番1号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 中野 幹生 (NAKANO, Mikio) [JP/JP]; 〒3510114 埼玉県和光市

- (74) 代理人: 特許業務法人オカダ・フシミ・ヒラノ (OKADA, FUSHIMI AND HIRANO, PC); 〒1020074 東京都千代田区九段南3丁目2番7号、NE九段ビル 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, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY,

[続葉有]

(54) Title: DEVICE, METHOD, AND PROGRAM FOR PERFORMING INTERACTION BETWEEN USER AND MACHINE

(54) 発明の名称: ユーザと機械とのインタラクションを実施するための装置、方法、およびプログラム



- AA ... DOMAIN a
- BB ... DOMAIN b
- CC ... DOMAIN n
- 11 ... VOICE RECOGNITION UNIT
- 13 ... DOMAIN SELECTION UNIT
- 31a ... VOICE COMPREHENSION UNIT
- 33a ... VOICE KNOWLEDGE
- 35a ... STATE STORAGE UNIT
- 37a ... BEHAVIOR EXTRACTION UNIT
- 39a ... BEHAVIOR KNOWLEDGE
- 41a ... TASK EXTRACTION UNIT
- 43a ... TASK KNOWLEDGE
- 25 ... SUB TASK KNOWLEDGE
- 23 ... TASK PLANNING UNIT
- 27 ... TASK PLAN STORAGE UNIT
- 17 ... BEHAVIOR SELECTION UNIT
- 19 ... BEHAVIOR EXECUTION UNIT
- 15 ... DOMAIN STORAGE UNIT
- 21 ... BEHAVIOR HISTORY STORAGE UNIT

(57) Abstract: There is provided a device for performing interaction between a user and a machine. The device includes a plurality of domains corresponding to a plurality of stages in the interaction. Each of the domains has voice comprehension means which understands the content of the user's voice. The device includes: means for recognizing the user's voice; means for selecting a domain enabling the best voice comprehension result as the domain; means for referencing task knowledge of the domain and extracting a task

[続葉有]

WO 2007/013521 A1



KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

添付公開書類:

— 国際調査報告書

correlated to the voice comprehension result; means for obtaining a sub task sequence correlated to the extracted task; means for setting the first sub task of the sub task sequence as the sub task and updating the domain to which the sub task belongs as the domain; means for extracting a behavior or sub task end flag correlated to the voice comprehension result and the subtask; and means for causing the machine to execute the extracted behavior.

(57) 要約: 本発明は、ユーザと機械とのインタラクションを実施する装置を提供する。この装置は、インタラクションにおける複数の段階に対応する複数のドメインを有し、各ドメインは、ユーザの音声の内容を理解して音声理解結果を出力する音声理解手段を備えている。この装置は、ユーザの音声を認識する手段と、最良の音声理解結果をとるドメインを当該ドメインとして選択する手段と、当該ドメインのタスク知識を参照して音声理解結果に関連付けられたタスクを抽出する手段と、抽出されたタスクに関連付けられたサブタスク系列を求める手段と、サブタスク系列の先頭のサブタスクを当該サブタスクとし、当該サブタスクの属するドメインを当該ドメインとして更新する手段と、当該ドメインの行動知識を参照して、音声理解結果およびサブタスクに関連付けられた行動またはサブタスク終了フラグを抽出する手段と、抽出された行動を機械に実行させる手段と、を備える。

明 細 書

ユーザと機械とのインタラクションを実施するための装置、方法、およびプログラム

技術分野

[0001] 本発明は、ユーザと機械とのインタラクションを実施するための装置、方法、およびプログラムに関する。

背景技術

[0002] 従来、ロボットまたはカー・ナビゲーションなどのような機械と、人間とのインターフェースとして、音声対話インターフェースが開発されている。対話インターフェースを有する機械は、ユーザとの対話や、対話で得られた情報に基づいて実行される自身の行動などによって、ユーザとのインタラクションを実施する。

[0003] 例えば、特開2003-316385号公報には、ユーザと機械が自由に対話するために、ユーザの発話を音声認識した結果に対応したタスク(ドメイン)を複数のタスクより選んで対話シーケンスを自動生成し、さらにユーザが発話した場合は必要に応じて対応したタスクの対話シーケンスに遷移し、もとの対話シーケンスに復帰する技術が開示されている。

発明の開示

発明が解決しようとする課題

[0004] しかしながら、従来技術では、ユーザ発話の音声認識結果に応じてタスク(ドメイン)を選択するので、車内の雑音などによって誤った認識が行われたときなどに、想定外のタスク(ドメイン)に遷移して対話の文脈が齟齬をきたし、ユーザが戸惑う事態が考えられる。

[0005] 本発明の目的は、対話の文脈を保って複数のドメインを遷移するようにユーザと機械とのインタラクションを実施することが可能な手法を提供することである。

課題を解決するための手段

[0006] 本発明は、ユーザと機械とのインタラクションを実施する装置を提供する。この装置は、インタラクションにおける複数の段階に対応する複数のドメインを有し、該複数の

ドメインのそれぞれは、ユーザの音声の内容を理解して音声理解結果を出力する音声理解手段を備えている。この装置は、マイクロフォンで検出された信号からユーザの音声を認識する手段と、音声理解手段のそれぞれにユーザの音声を送り、音声理解手段のそれぞれから音声理解結果を受け取り、最良の音声理解結果をとる音声理解手段が含まれるドメインを当該ドメインとして選択する手段と、複数のドメインのそれぞれに含まれるタスク知識のうち当該ドメインのタスク知識を参照して、音声理解結果に関連付けられたタスクを抽出する手段と、タスクの種類に関連付けられる複数のサブタスクを含むサブタスク知識を参照して、前記抽出されたタスクに関連付けられたサブタスク系列を求める手段と、サブタスク系列の先頭のサブタスクを当該サブタスクとし、当該サブタスクの属するドメインを当該ドメインとして更新する手段と、複数のドメインのそれぞれに含まれる行動知識のうち当該ドメインの行動知識を参照して、音声理解結果および前記サブタスクに関連付けられた行動またはサブタスク終了フラグを抽出する手段と、抽出された行動を機械に実行させる手段と、を備える。

- [0007] この発明により、複数のドメインに跨る人と機械のインタラクションを滑らかに実行できる。
- [0008] 本発明の一実施形態によると、サブタスク知識は、タスクに関連付けられる1つ以上のサブタスクに関する知識と、サブタスクに関連付けられるドメインに関する知識と、を含む。これにより、タスクを達成する間、ドメインの文脈を保持することができる。
- [0009] 本発明の一実施形態によると、音声理解手段のそれぞれは、対応するドメインに関連性の高い複数の文パターンを含む音声知識を参照して、音声と複数の文パターンのそれぞれとの適合度を算出し、適合度が最も高い文パターンを選択し、選択された文パターンおよび該文パターンの適合度を音声理解結果として出力する。これにより、ドメインの文脈を保つことができる。
- [0010] 本発明の一実施形態によると、選択する手段は、適合度に複数のドメインごとに設定された重みを乗じて信頼度を算出し、最大の信頼度をとるドメインを当該ドメインとして選択する。これにより、ドメインの文脈を保つことができる。
- [0011] 本発明の一実施形態によると、更新する手段は、行動またはサブタスク終了フラグを抽出する手段がサブタスク終了フラグを抽出すると、サブタスク系列における当該

サブタスクの次のサブタスクを新たな当該サブタスクとして更新し、新たな当該サブタスクの属するドメインを当該ドメインとして更新する。これによりタスクを円滑に達成することができる。

[0012] さらに、本発明は、ユーザと機械とのインタラクションを実施するための方法を提供する。この方法は、マイクロフォンで検出された信号からユーザの音声を認識するステップと、インタラクションにおける複数の段階に対応する複数のドメインのそれぞれにユーザの音声を送るステップと、複数のドメインのそれぞれにおいて、音声の内容を理解して音声理解結果を出力するステップと、複数のドメインのそれぞれから音声理解結果を受け取るステップと、複数の音声理解結果のうち最良の音声理解結果をとるドメインを、当該ドメインとして選択するステップと、複数のドメインのそれぞれに含まれるタスク知識のうち当該ドメインのタスク知識を参照して、音声理解結果に関連付けられたタスクを抽出するステップと、タスクの種類に関連付けられた複数のサブタスクを含むサブタスク知識を参照して、抽出されたタスクに関連付けられたサブタスク系列を求めるステップと、サブタスク系列の先頭のサブタスクを当該サブタスクとし、当該サブタスクの属するドメインを当該ドメインとして更新するステップと、複数のドメインのそれぞれに含まれる行動知識のうち当該ドメインの行動知識を参照して、音声理解結果およびサブタスクに関連付けられた行動またはサブタスク終了フラグを抽出するステップと、抽出された行動を機械に実行させるステップと、を含む。

[0013] さらに、本発明は、ユーザと機械とのインタラクションを実施するためのプログラムを提供する。このプログラムは、マイクロフォンで検出された信号からユーザの音声を認識する機能と、ユーザとのインタラクションにおける複数の段階に対応する複数のドメインのそれぞれにユーザの音声を送る機能と、複数のドメインのそれぞれにおいて、音声の内容を理解して音声理解結果を出力する機能と、複数のドメインのそれぞれから、音声理解結果を受け取る機能と、複数の音声理解結果のうち最良の音声理解結果をとるドメインを、当該ドメインとして選択する機能と、複数のドメインのそれぞれに含まれるタスク知識のうち当該ドメインのタスク知識を参照して、音声理解結果に関連付けられたタスクを抽出する機能と、タスクの種類に関連付けられる複数のサブタスクを含むサブタスク知識を参照して、抽出されたタスクに関連付けられたサブタスク

系列を求める機能と、サブタスク系列の先頭のサブタスクを当該サブタスクとし、当該サブタスクの属するドメインを当該ドメインとして更新する機能と、複数のドメインのそれぞれに含まれる行動知識のうち当該ドメインの行動知識を参照して、音声理解結果およびサブタスクに関連付けられた行動またはサブタスク終了フラグを抽出する機能と、抽出された行動を機械に実行させる機能と、をコンピュータに実現させる。このプログラムは、コンピュータ読取り可能な記録媒体に記録される。

図面の簡単な説明

- [0014] [図1]音声インタラクティブ装置の機能ブロック図である。
[図2]発話理解処理のフローチャートである。
[図3]行動選択処理のフローチャートである。
[図4]ステップS209の行動実行・タスク抽出処理のフローチャートである。

符号の説明

- [0015] 10 音声インタラクティブ装置
11 音声認識部
13 ドメイン選択部
17 行動選択部
19 行動実行部
23 タスク計画部
25 サブタスク知識
31a～31n 音声理解部
37a～37n 行動抽出部
41a～41n タスク抽出部
43a～43n タスク知識

発明を実施するための最良の形態

[0016] 1. 概略

次に図面を参照して、この発明の実施の形態を説明する。図1は、本発明の一実施形態による音声インタラクティブ装置10の機能ブロック図である。

- [0017] 音声インタラクティブ装置10は、例えば、種々の演算を実行するCPU、演算結果を

一時記憶するためのメモリ、学習結果等を記憶するROM、データの入出力を行うインターフェース等を備えるマイクロコンピュータにより実現される。図1に示された音声インタラクティブ装置10の各機能ブロックの一部または全部は、ソフトウェア、ファームウェア、あるいはハードウェアの何れでも実現することができる。

- [0018] 音声インタラクティブ装置10は、人間と機械とのインタラクションを実施するための装置である。音声インタラクティブ装置10は、機械に備えられたマイクロフォンなどの音声検知手段(図示せず)から人間の音声を受け取り、機械の発声手段または移動手段など(図示せず)へ動作指令を送信する。
- [0019] ここで、本発明における「機械」とは、ロボット、カー・ナビゲーション、電化製品など、インターフェースに音声を用いるもののことをいう。
- [0020] 音声インタラクティブ装置10は、機械の内部に組み込む形式でも、機械の外部から有線または無線で機械と通信する形式でも良い。
- [0021] 音声インタラクティブ装置10は、複数のドメインa~nを備える。
- [0022] ここで、本発明における「ドメイン」とは、人間と機械との間のインタラクションを構成する複数の段階(フェーズ)のことをいう。
- [0023] 本実施形態では、ドメインa~nは、人間の要求の種類を理解するために対話を行う「要求理解ドメイン」と、人間の要求に応えるために行動を実行する「非要求理解ドメイン」とに大別される。
- [0024] 「要求理解ドメイン」は、人間の発話内容を推定したり、不明な点を質問するなどの対話を通して、例えば「人を呼びに行く」というような人間からの要求を理解する。要求理解ドメインは、「人を呼びに行く」の他にも、要求の内容に応じて別々なドメインが用意されており、これら複数の要求理解ドメインの中からユーザの発話内容に最適な1つが選択される。ドメインの選択は、後述する発話理解処理によって行われる。1つの要求理解ドメインを選択することにより、ユーザの要求の内容が理解されたことになる。
- [0025] 「非要求理解ドメイン」は、例えば「目標位置の情報を得る」、「目標位置へ移動する」、「メッセージを伝える」というような行動の内容に応じて別々なドメインが用意されている。

[0026] 複数のドメインa～nは、それぞれ、音声理解部31a～n、音声知識33a～n、状態格納部35a～n、行動抽出部37a～n、行動知識39a～n、タスク抽出部41a～n、タスク知識43a～nを備える。音声知識、行動知識、およびタスク知識は、ドメインの種類に応じて関連性の高い文パターン、行動、およびタスクがそれぞれ記憶されている。

[0027] 本実施形態では、階層的に「タスク」、「サブタスク」、および「行動」という言葉を使い分ける。はじめに、ドメイン選択部13が、ユーザの発話内容に最適な要求理解ドメインを選択して、選択されたドメイン(a～nのいずれか) (以下「当該ドメイン」という)のタスク抽出部(41a～41nのいずれか)が、ユーザの要求内容を満足するためにロボットが行うべき「タスク」を抽出する。次に、タスク計画部23が、抽出された「タスク」を行うための「サブタスク」の系列を決定する。そして、行動選択部17が、「サブタスク」系列にしたがって、「サブタスク」を行うための「行動」を選択する。つまり、「タスク」、「サブタスク」、「行動」の内容は、タスク→サブタスク→行動の順で細分化される。

[0028] 音声インタラクティブ装置10で行われる処理は、「発話理解処理」および「行動選択処理」に分けられる。「発話理解処理」は、主に、音声認識部11、ドメイン選択部13、各ドメインの音声理解部31a～n、当該ドメイン格納部15によって実施される。「行動選択処理」は、主に、行動選択部17、行動実行部19、行動履歴格納部21、当該ドメイン格納部15、当該ドメインの行動抽出部(37a～37nのいずれか)、タスク抽出部(41a～41nのいずれか)、タスク計画部23によって実施される。以下に、「発話理解処理」および「行動選択処理」の詳細について説明する。

[0029] 2. 発話理解処理

次に、図2を参照して音声インタラクティブ装置10の発話理解処理について説明する。図2は、発話理解処理のフローチャートである。

[0030] ステップS101において、マイクロフォンなどの検知手段(図示せず)からユーザの発話が入力されると、音声認識部11が音声を認識する。音声認識部11は、隠れマルコフモデルなど公知の手法を用いて、単語列および各単語の認識結果の信頼度の形で音声認識結果を出力する。信頼度は、一般に0～1の値をとる。単語列に含まれる単語のそれぞれは、複数の候補を有していても良い。雑音のある環境にいる場

合は、ビームフォーミングなどの公知の手法を用いて、マイクロフォンからの入力信号から音声を抽出して、抽出された音声を認識しても良い。

[0031] ステップS103において、ドメイン選択部13が、音声認識結果を各ドメインa～nの音声理解部31a～31nへ送信する。また、ドメイン選択部13は、音声認識結果を音声認識部11から受け取った時刻から所定の期間(たとえば3秒間)、当該ドメイン選択部15に発話終了フラグを立てる。

[0032] ステップS105において、各ドメインの音声理解部31a～31nが、ドメイン選択部13より入力された音声認識結果に含まれる単語列と、ドメインに応じて用意された音声知識33a～33nにある文パターンとの「適合度」を算出する。

[0033] 音声知識33a～33nは、音声認識結果を理解するためにドメインに応じて用意された複数の「文パターン」から構成される。「文パターン」は、例えば、

「<日にち>の<場所>の<情報種別>を教えて」

というものがある。この例示した文パターンでは、<日にち>、<場所>、<情報種別>という3つのカテゴリを含んでいる。各カテゴリは対応したキーワードリストを持つ。

[0034] 音声認識結果に含まれる単語のそれぞれは、文パターンのカテゴリ別に用意されたキーワードと一致すると、そのカテゴリに適合される。「適合度」は、例えばある文パターンが含むカテゴリのうち、単語を適合されたカテゴリの数や割合である。

[0035] 例えば、「明日の天気を教えて」という音声認識結果を、

「<日にち>の<場所>の<情報種別>を教えて」

という文パターンに適合させると、

<日にち>=明日、<場所>="空白"、<情報種別>=天気

という結果となる。このとき、3つのカテゴリのうち2つのカテゴリが音声認識結果と適合しているので、適合度は「2」または「2/3」と表される。

[0036] ステップS107において、音声理解部31a～31nが、最も適合度の高い文パターンを選択する。そして、選択した文パターンの各カテゴリに適合する単語を当てはめた文(以下「適合済み文パターン」という)と、この文パターンの適合度とを、「音声理解結果」として状態格納部35a～35nに格納する。

[0037] 「適合済み文パターン」は、上記の例に従うと、

「<日にち=明日>の<場所=“空白”>の

<情報種別=天気>を教えて」

のように表される。また、「適合済み文パターン」に各単語の信頼度を加えても良い。

例えば、「明日」という単語の音声認識の信頼度を0.9、「天気」という単語の音声認識の信頼度を0.7とすると、適合済み文パターンは、

「<日にち=明日(0.9)>の<場所=“空白”(0.0)>の

<情報種別=天気(0.7)>を教えて」

のように表される。

- [0038] ステップS109において、音声理解部31a～31nは、選択した文パターンの適合度をドメイン選択部13に送る。
- [0039] ステップS111において、ドメイン選択部13が、各ドメインの音声理解部31a～31nから受け取った適合度から「選択信頼度」を算出する。
- [0040] 「選択信頼度」は、ドメイン別に設定された「重み」を適合度に乗じて算出される。
- [0041] 「重み」は、当該ドメイン格納部15に格納されている当該ドメインに対して1より大きい値が設定され、他のドメインに対しては1が設定される。これは、当該ドメインの対話や行動は連続して行われる可能性が高いと仮定しているためである。当該ドメインとは、直前のユーザ発話またはサブタスクの属するドメインである。
- [0042] 当該ドメインのない場合は、全ての重みが等しい値(たとえば1)に設定される。
- [0043] 当該ドメインの重みは、要求理解ドメインであれば他ドメインより少し大きい値、非要求理解ドメインであれば大きな値としても良い。このように重み付けると、要求選択ドメインは固定されにくく、それ以外のドメインは固定されやすくなる。これは、ユーザの要求を理解するには対話の試行錯誤が求められることが多いが、ひとたび要求を理解すると、その後の対話や行動の文脈は固定されやすい、と仮定しているためである。
- [0044] ステップS113において、ドメイン選択部13は、選択信頼度が最も大きいドメインを当該ドメインとして選択し、この選択したドメインの情報を当該ドメイン格納部15に格納する。複数のドメイン選択信頼度が同じだった場合は、当該ドメインを選択する。このようにドメインを推定すると、当該ドメインが選ばれやすくなる。選択信頼度を算出

せずに、適合度が最も大きいドメインを当該ドメインとしても良い。

[0045] 以上に、音声インタラクティブ装置10の音声理解処理について説明した。

[0046] 入力されたユーザ発話に対して以上の発話理解処理を行った後、当該ドメイン格納部15には、現在のシステムが行っている対話・行動のドメインを表す「当該ドメイン」と、ユーザの発話の理解が終了したことを示す「発話終了フラグ」が格納される。

[0047] また、各ドメインの状態格納部35a～35nには、最も適合度の高かった文パターンの適合済み文パターンおよび適合度を含む「音声理解結果」が格納される。

[0048] 3. 行動選択処理

次に、図3および図4を参照して音声インタラクティブ装置10の行動選択処理について説明する。図3は、行動選択処理のフローチャートである。

[0049] 図3に示す行動選択処理は、行動選択部17によって、所定の周期(たとえば100ミリ秒)で実行される。

[0050] ステップS201において、行動選択部17は、行動履歴格納部21を参照して、ロボットが行動を実行しているかどうかを確認する。行動履歴格納部21は、「直前に実行された行動」、「その行動の開始時刻」、「行動の終了時刻」を記録している。行動が実行中の場合終了時刻は記録されていないので、行動選択部17は、行動履歴格納部21の「行動の終了時刻」の有無を確認することにより、行動が実行中かどうかを把握する。行動が実行中でない場合、ステップS203に進む。行動が実行中の場合、今回の周期の処理は終了し、次の周期まで待機する。

[0051] ステップS203において、行動選択部17は、ロボットが行動を終了した直後かどうかを確認する。この処理は、たとえば、行動履歴格納部21の「行動の終了時刻」を参照して、現在の時刻が「行動の終了時刻」から所定時間経過しているかを確認する。行動が終了した直後ではない場合、ステップS205に進む。行動が終了した直後だった場合、ステップS207に進む。

[0052] ステップS205において、行動選択部17は、当該ドメイン格納部15の発話終了フラグを参照して、ユーザの発話が終了しているかどうかを確認する。発話が終了している場合、ステップS209に進み、「行動選択・タスク抽出」処理を行う。発話が終了していない場合、今回の周期の処理は終了し、次の周期まで待機する。

- [0053] ステップS207において、行動選択部17は、行動履歴格納部21の「直前に実行された行動」を参照して、直前の行動が「ユーザへの質問」かどうかを確認する。直前の行動が「ユーザへの質問」ではない場合、ステップS209に進み、「行動選択・タスク抽出」処理を行う。直前の行動が「ユーザへの質問」だった場合、今回の周期の処理は終了し、次の周期まで待機する。
- [0054] 図4は、ステップS209の行動実行・タスク抽出処理のフローチャートである。
- [0055] なお、説明の便宜上、以下の処理では当該ドメインとしてドメインaが選択されている場合を扱っている。実際の処理では、当該ドメインは、ドメインa～nの中から適宜選択される。
- [0056] 行動実行・タスク抽出処理では以下のステップが実行される。
- [0057] (1)行動選択部17が、当該ドメイン格納部15を参照して当該ドメインaを調べ、当該ドメインの行動抽出部37aに行動選択を要求する(ステップS301)。
- [0058] (2)行動選択部17より行動選択が要求されると、当該ドメインaの行動抽出部37aが以下の処理を行う。
- [0059] (2-1)当該ドメインaが理解要求ドメインの場合(ステップS303のYES)、
a)行動知識39aの中の行動選択規則を参照して、状態格納部35aの音声理解結果に対応した行動またはサブタスク終了フラグを選択する(ステップS305)。
- [0060] ここで、行動選択規則とは、状態格納部35a～nに格納されている音声理解結果、当該サブタスク、及び行動履歴のいずれか又は全てと、行動またはサブタスク終了フラグとの関係を表したものであり、音声対話の研究分野においてアクションセレクションを行うために一般的に用いられる規則である。
- [0061] 行動選択規則は、例えば、「音声理解結果で文パターンに”空白”のある場合に、空白を問う行動を生成する」、「音声理解結果が良い場合(適合状態が良く適合度が高い)、サブタスク終了フラグを選択する」、「音声理解結果が良くなく、行動履歴にある直前の行動がユーザに対する質問の場合は、繰り返し要求発話生成をする」、「音声理解結果が良くなく、行動履歴にある直前の行動がユーザへの了解発話の場合は、再び発話を問う依頼発話生成」などである。
- [0062] また、行動知識には、「言語生成の知識」も含まれる。選択された行動が上記の「空

白を問う行動を生成する」のような発話に関するものである場合、行動と発話内容の文章とを対応付ける。

[0063] b) 行動が選択された場合、行動選択部17に行動を出力する(ステップS315)。

[0064] c) サブタスク終了フラグが選択された場合、行動抽出部37aからタスク抽出部41aへタスク選択が要求される(ステップS309)。

[0065] d) タスク抽出部41aが、タスク知識43aの中のタスク抽出規則を参照して音声理解結果に対応したタスクを、タスク計画部23に出力する(ステップS311)。

[0066] ここで、タスク抽出規則とは、状態格納部35a～nに格納されている音声理解結果と、タスク計画部23に出力すべきタスクとの関係を表す規則である。例えば、「人を呼ぶ要求理解のドメインa」には、「<人>を呼んできて」という文パターンの<人>カテゴリに単語が適合されている適合済み文パターンを含む音声理解結果に対して、<タスク=人を呼びに行く、人=<人>>という出力タスクが関連付けられている。

[0067] e) タスク計画部23は、サブタスク知識25を利用してタスクをサブタスク系列に変換し、タスク計画格納部27に格納する(ステップS313)。

[0068] ここで、サブタスク知識25は、タスクをサブタスクに分解するための知識(規則)を持ち、サブタスクはそれぞれドメインと関連付けられている。

[0069] タスクのサブタスクへの分解は、例えば以下のように行われる。

<タスク=呼ぶ、人=<X>>

というタスクは、

(i) <サブタスク=人の位置を知る、人=X>

(ii) <サブタスク=人に近づく、人=X>

(iii) <サブタスク=呼ばれていることを伝える、人=X>

というサブタスクの系列に分解される。

[0070] また、各サブタスクは、ドメインと対応付けられており、例えば サブタスク(i)が「人の位置を知る情報取得ドメイン」と対応し、サブタスク(ii)が「移動するドメイン」と対応し、サブタスク(iii)が「呼ばれているという情報提供ドメイン」と対応するよう設定されている。

[0071] (2-2) 当該ドメインが非理解要求ドメインの場合(ステップS303のNO)、行動知識

39aの中の行動選択規則(上記(2-1)のa)に詳述)を参照して、状態格納部35aの音声理解結果、当該サブタスク及び行動履歴に対応した行動またはサブタスク終了フラグを選択し、行動選択部17に出力する(ステップS317)。

[0072] (3) 行動抽出部37aより行動またはサブタスク終了フラグが入力されると(ステップS319)、行動選択部17が、

(3-1) 行動が入力された場合は、その行動を行動実行部19に出力する(S329)。行動実行部19は、行動選択部17より行動が入力されるとそれを実行し、実行開始時間を行動履歴格納部21に格納する。実行を終了すると、当該行動と実行終了時間を行動履歴格納部21に格納する。

[0073] (3-2) サブタスク終了フラグが入力された場合は、当該ドメイン格納部15の当該ドメインをクリアし、タスク計画部23にサブタスクを要求する(ステップS321)。

[0074] a) タスク計画部23が、タスク計画格納部27を参照して、サブタスク系列がある場合は先頭のサブタスクを行動選択部17に出力し、サブタスク系列がない場合はタスク無しフラグを行動選択部17に出力する(ステップS323)。

[0075] b) サブタスクが入力された場合は、そのサブタスクのドメイン(a~nのいずれか)を当該ドメインとして当該ドメイン格納部15に格納し、当該ドメインの行動抽出部(37a~37nのいずれか)にそのサブタスクを出力する。当該ドメインの行動抽出部(37a~37nのいずれか)は、当該サブタスクを状態格納部(35a~35nのいずれか)に格納する(ステップS327)。

[0076] c) タスク無しフラグが送られてきた場合は、全ての格納部のデータをクリアする(ステップS331)。

[0077] 以上に図3および図4を参照して、行動選択処理について説明した。

[0078] 行動選択処理が行われると、各ドメインの状態格納部35a~35nには、最も適合度の高かった文パターンの適合済み文パターンおよび適合度を含む「音声理解結果」に加えて「サブタスク」が格納される。

[0079] タスク計画格納部27には、サブタスク系列が格納される。

[0080] 行動履歴格納部21には、「直前の行動」、「行動の実行終了時間」が格納される。

[0081] また、当該ドメイン格納部15の当該ドメインが更新される場合がある。

[0082] 4. 実施例

4.1 概略

次に、4つのドメインa～dを持つロボットとユーザ(アンリさん)との対話を通じたインタラクションの実施例を示す。

[0083] この実施例でロボットが備える4つのドメインa～dは、「人を呼びに行く要求理解ドメインa」、「人の位置を知る情報取得ドメインb」、「移動するドメインc」、「呼ばれているという情報提供ドメインd」である。

[0084] ロボットとユーザが次のような対話を行う場合を考える。

[0085] アンリさん 「ジダンさんを呼んできて」
 ロボット 「ジダンさんはどこにいますか？」
 アンリさん 「ドレッシングルーム」
 ロボット 「ジダンさんはどこにいますか？」
 アンリさん 「ドレッシングルーム」
 ロボット 「わかりました」
 (ジダンさんのところに移動を始める)
 ロボット 「アンリさんが呼んでいます」

[0086] このような対話の間に音声インタラクティブ装置10で行われる処理について以下に説明する。

[0087] 4.2 発話理解処理

初期状態では、全ての格納部の中身は空である。

[0088] マイクロフォンよりアンリさんの発話「ジダンさんを呼んできて」が入力されると、音声認識部11は、「ジダン」、「さん」、「を」、「呼んで」、「きて」という単語列と認識し、各単語の信頼度(0～1)と合わせてドメイン選択部13に出力する。ここでは良好な音声認識が行われ、各単語はそれぞれ高い信頼度で認識された。

[0089] ドメイン選択部13は、単語列を各ドメインa～dの音声理解部31a～31dへ送信し、当該ドメイン格納部15に所定の期間、発話終了フラグを立てる。

[0090] 音声理解部31a～31dのそれぞれは、音声知識33a～33dにある文パターンにこの単語列を適合させる。

- [0091] 例えば「人を呼びに行く要求理解ドメインa」には、「<人>を<呼んで>きて」という文パターンがあるので、この文パターンに適合させてみる。すると、<人>カテゴリが有するキーワードリストと”ジダン”、”さん”が適合するため、「人＝ジダンさん」が入力され、また、<呼んで>が一致する。
- [0092] この場合、文パターンが含む全て(2つ)のカテゴリに単語が適合しており、また”ジダン”、”さん”、”呼んで”の信頼度が高かったので、高い適合度が算出される。
- [0093] 「人を呼びに行く要求理解ドメインa」の他の文パターンに適合させた結果、この文パターンの適合度が最も高かったので、適合済み文パターン「<人＝ジダンさん>を<呼んで>きて」と適合度とを音声認識結果として状態格納部35aに格納し、適合度をドメイン選択部13に出力する。
- [0094] 他の3つのドメインb～dも同様に処理する。
- [0095] ドメイン選択部13は、各ドメインa～dの音声理解部31a～31dより入力された適合度よりドメイン選択信頼度を算出して比較し、最も大きいドメイン選択信頼度を有するドメインを当該ドメインとして当該ドメイン格納部15に格納する。
- [0096] この発話では、当該ドメイン格納部15に当該ドメインが記録されておらず、「人を呼びに行く要求理解ドメインa」の適合度が最大だったので、これを当該ドメインとする。
- [0097] 4.3 タスク抽出処理
行動選択部17は、100ミリ秒間隔で行動履歴格納部21を参照する。
- [0098] この時、行動履歴格納部21には何も格納されていないため、ロボットは行動実行中ではなく、また行動終了直後ではないと判断する。
- [0099] そして、当該ドメイン格納部15の発話終了フラグを検出して発話終了と判断する。そこで行動選択部17は行動選択処理を行う。
- [0100] 行動選択処理では、行動選択部17が当該ドメイン格納部15を参照し、当該ドメインである「人を呼びに行く要求理解ドメインa」の行動抽出部37aに行動選択を要求する。
- [0101] 当該ドメインは要求理解ドメインなので、行動抽出部37aは、状態格納部35aにある音声認識結果と、行動知識39aにある行動抽出規則を照らし合わせる。
- [0102] ここでは音声理解結果が良好なので、サブタスク終了フラグが選択される。

[0103] そこでタスク知識43aのタスク抽出規則と照らし合わせ、〈タスク=呼ぶ、人=〈人〉〉という規則を選択し、〈タスク=呼ぶ、人=ジダンさん〉というタスクを得て、これをタスク計画部23に出力する。

[0104] タスク計画部23は、上記タスクを受け、サブタスク知識25を参照して以下のサブタスク系列に分解してタスク計画格納部27に格納し、先頭のサブタスク(i)を行動選択部17に出力する。

[0105] (i) 〈サブタスク=人の位置を知る、人=ジダンさん〉…………ドメインb

(ii) 〈サブタスク=人に近づく、人=ジダンさん〉…………ドメインc

(iii) 〈サブタスク=呼ばれていることを伝える、人=ジダンさん〉…ドメインd

行動選択部17は、サブタスク(i)のドメインである「人の位置を知る情報取得ドメインb」を当該ドメイン格納部15に格納し、当該ドメインの行動抽出部37bにサブタスク(i)を出力する。

[0106] 当該ドメインの行動抽出部37bは、入力されたサブタスク(i)を状態格納部37bに格納する。

[0107] 4.4 行動選択処理1

行動選択部17は、再度(100ミリ秒後)、行動履歴格納部21と当該ドメイン格納部15を参照してユーザ発話終了フラグを検出し、行動選択処理を行う。

[0108] 行動選択部17は、当該ドメイン格納部15を参照して当該ドメイン「人の位置を知る情報取得ドメインb」を得、当該ドメインの行動抽出部37bに行動選択を要求する。

[0109] 当該ドメインの行動抽出部37bは、行動知識39bにある行動選択規則を参照して音声理解結果とサブタスク(i)に対応する行動を選択する。

[0110] 当該ドメインの状態格納部35bには

「〈人=ジダンさん〉はどこ=“空白”」

という音声理解結果が格納されており、サブタスクが

(i)〈サブタスク=人の位置を知る、人=ジダンさん〉

なので、

〈行動=位置を尋ねる、人=〈人〉〉

という規則を選択し、

<行動＝位置を尋ねる、人＝ジダンさん>

という行動を得る。

[0111] この行動を、行動知識39bにある言語生成の知識を用いて

「ジダンさんはどこにいますか？」

という文に変換し、行動選択部17に出力する。

[0112] 行動選択部17はこの行動(文)を行動実行部19に出力する。

[0113] 行動実行部19は、「ジダンさんはどこにいますか？」と発話し、ユーザへの質問という当該行動と実行終了時間を行動履歴格納部21に格納する。

[0114] ユーザが「ドレッシングルーム」と応答発話をする、音声認識部11は “ジダンさん”と低い評価値で認識する。

[0115] ドメイン選択部13はこの認識結果を各ドメインの音声理解部31a～31dに送信し、当該ドメイン格納部15に発話終了フラグを立てる。

[0116] この発話は評価値が低いため、全ドメインで低い適合度が算出される。

[0117] 適合度が低いため、全ドメインの音声理解部は誤認識を検出し、適合度と合わせて状態格納部35a～35dに格納し、適合度をドメイン選択部13に出力する。

[0118] ドメイン選択部13は、当該ドメイン格納部15に格納されている当該ドメインを参照し、「人の位置を知る情報取得ドメインb」であることを知り、このドメインに高い重みをつけて、ドメイン選択信頼度を計算する。その結果、再び「人の位置を知る情報取得ドメインb」が選択されたため、当該ドメイン格納部15の当該ドメインを更新しない。

[0119] 行動選択部17は、行動終了直後でユーザ発話終了であることを知り、行動選択処理を行う。

[0120] 行動選択部17は当該ドメイン格納部15を参照し、当該ドメインを「人の位置を知る情報取得ドメインb」と知り、当該ドメインの行動抽出部37bに行動選択を要求する。

[0121] 当該ドメインの行動抽出部37bは、状態格納部35bに格納されている「誤認識」と低い適合度という音声理解結果と、行動選択規則を参照し、

<行動＝繰り返し要求発話生成、相手＝ユーザ>

という行動を得る。

[0122] さらに、行動抽出部37bは、行動知識39bにある言語生成の知識を参照して、この

行動を、

「ジダンさんはどこにいますか？」

という文に変換して行動選択部17に出力する。

[0123] 行動選択部17はこれを行動実行部19に出力する。

[0124] 行動実行部19は、「もう一度お願いします」と発話し、ユーザへのお願い(再質問)という当該行動と実行終了時間を行動履歴格納部21に格納する。

[0125] ユーザが再度、

「ドレッシングルーム」

と発話すると、今度は高い評価値で認識される。

[0126] 同様の処理が行われ、当該ドメイン格納部15に発話終了フラグが立ち、状態格納部35bに

「<人=ジダンさん>は<どこ=ドレッシングルーム>」

という適合済み文パターンが、高い適合度で音声理解結果として格納される。

[0127] 当該ドメインbの適合度が最も高いので、ドメイン選択部13は当該ドメイン格納部15の当該ドメインを更新しない。

[0128] 行動選択部17は、当該ドメイン格納部15の発話終了フラグを検出して発話終了と判断し、行動選択処理を行う。

[0129] 行動選択部17は、当該ドメインが「人の位置を知る情報取得ドメインb」であることを知り、当該ドメインの行動抽出部37bに行動選択を要求する。

[0130] 行動抽出部37bは、

「<人=ジダンさん>は<どこ=ドレッシングルーム>」

という良好な音声理解結果と、

(i)<サブタスク=人の位置を知る、人=ジダンさん>

というサブタスクを行動選択規則と参照し、文パターンの全カテゴリに単語が適合しかつ適合度が高いことから、サブタスク終了フラグを行動選択部17に出力する。

[0131] サブタスク終了フラグを受け、行動選択部17は当該ドメイン格納部15の当該ドメイン「人の位置を知る情報取得ドメインb」をクリアし、タスク計画部23にサブタスクを要求する。

- [0132] タスク計画部23は、タスク計画格納部27を参照して
(ii)〈サブタスク=人に近づく、人=ジダンさん〉「移動するドメインc」
というサブタスクを行動選択部17に出力する。
- [0133] 行動選択部17は、「移動するドメインc」を当該ドメインとして当該ドメイン格納部15
に格納し、当該ドメインの行動抽出部37cにサブタスク(ii)を出力する。
- [0134] 当該ドメインの行動抽出部37cは、状態格納部35cにサブタスク(ii)を格納する。
- [0135] 4.5 行動選択処理2
行動選択部17は再度(100ミリ秒後)、行動履歴格納部21と当該ドメイン格納部15
を参照してまだ立っている発話終了フラグを検出し、行動選択処理を行う。
- [0136] 行動選択部17は、当該ドメインである「移動するドメインc」の行動抽出部37cに行
動選択を要求する。
- [0137] 当該ドメインの行動抽出部37cは、状態格納部35cにある、
「〈人=ジダンさん〉は〈どこ=ドレッシングルーム〉」
という音声理解結果と、
(ii)〈サブタスク=人に近づく、人=ジダンさん〉
というサブタスクを行動選択規則と参照し、
〈行動=了解発話生成後移動、位置=〈どこ〉〉
という規則を選択して、
〈行動=了解発話生成後移動、位置=ドレッシングルーム〉
という行動を得、行動選択部17に出力する。
- [0138] 行動選択部17は行動実行部19にこれを出力する。
- [0139] 行動実行部19は、
「分かりました」
と発話してからロボットをドレッシングルームに移動させ、到着したら移動という当該行
動と実行終了時間を行動履歴格納部21に格納する。
- [0140] 4.6 行動選択処理3
数分の後、行動選択部17は行動終了を検出し、行動選択処理を行う。
- [0141] 行動選択部17は、当該ドメインである「移動するドメインc」の行動抽出部37cに行

動選択を要求する。

- [0142] 当該ドメインの行動抽出部37cは、状態格納部35cにある、
(ii)〈サブタスク=人に近づく、人=ジダンさん〉
と行動履歴格納部21の行動履歴を行動選択規則と参照してサブタスク終了フラグを選択し、行動選択部17に出力する。
- [0143] 行動選択部17は当該ドメイン格納部15にある「移動するドメインc」をクリアし、タスク計画部23にサブタスクを要求する。
- [0144] タスク計画部23は、タスク計画格納部27を参照して
(iii)〈サブタスク=呼ばれていることを伝える、人=ジダンさん〉
「呼ばれているという情報の提供ドメインd」
というサブタスク(iii)を行動選択部17に出力する。
- [0145] 行動選択部17は「呼ばれているという情報提供ドメインd」を当該ドメイン格納部15に格納し、当該ドメインの行動抽出部37dに上記(iii)のサブタスクを出力する。
- [0146] 行動抽出部37dは状態格納部35dにサブタスク(iii)を格納する。
- [0147] 行動選択部17は、行動履歴格納部21を参照して行動実行直後であることを知り、行動選択処理を行う。
- [0148] 行動選択部17は、当該ドメインである「呼ばれているという情報の提供ドメインd」の行動抽出部37dに行動選択を要求する。
- [0149] 当該ドメインの行動抽出部37dは、状態格納部35dにある、
(iii)〈サブタスク=呼ばれていることを伝える、人=ジダンさん〉
を行動選択規則と照らし合わせ、
〈行動=人へ情報を伝える発話生成、人=ジダンさん〉
という行動を得る。
- [0150] さらに、行動抽出部37dは、行動知識39dにある言語生成の知識を参照して、この行動を、
「アンリさんが呼んでいます」
という発話文に変換して行動選択部17に出力する。
- [0151] 行動選択部17はこれを行動実行部19に出力する。

- [0152] 行動実行部19は、「アンリさんが呼んでいます」と発話し、情報を伝えるという当該行動と実行終了時間を行動履歴格納部21に格納する。
- [0153] 4.7 行動選択処理4
行動選択部17は、行動履歴格納部21を参照して行動終了を知り、行動選択処理を行う。
- [0154] 行動選択部17は、当該ドメインである「呼ばれているという情報の提供ドメインd」の行動抽出部37dに行動選択を要求する。
- [0155] 当該ドメインの行動抽出部37dは、状態格納部35dにある、
(iii)〈サブタスク＝呼ばれていることを伝える、人＝ジダンさん〉
と、行動履歴を行動選択規則と照らし合わせ、サブタスク終了フラグを選択し、行動選択部17に出力する。
- [0156] 行動選択部17は当該ドメイン格納部15にある「呼ばれているという情報の提供ドメインd」をクリアし、タスク計画部23にサブタスクを要求する。
- [0157] タスク計画部23はタスク計画格納部27を参照し、サブタスクがないのでタスク無しフラグを行動選択部17に出力する。
- [0158] 行動選択部17はタスク無しフラグを受け、全格納部のデータをクリアする。
- [0159] 以上にこの発明を特定の実施形態によって説明したが、この発明はこのような実施形態に限定されるものではない。

請求の範囲

- [1] ユーザと機械とのインタラクションを実施する装置であつて、
前記装置は、前記インタラクションにおける複数の段階に対応する複数のドメインを有し、該複数のドメインのそれぞれは、前記ユーザの音声の内容を理解して音声理解結果を出力する音声理解手段を備えており、
マイクロフォンで検出された信号から前記ユーザの音声を認識する手段と、
前記音声理解手段のそれぞれに前記ユーザの音声を送り、前記音声理解手段のそれぞれから前記音声理解結果を受け取り、最良の音声理解結果をとる音声理解手段が含まれるドメインを当該ドメインとして選択する手段と、
前記複数のドメインのそれぞれに含まれるタスク知識のうち前記当該ドメインのタスク知識を参照して、前記音声理解結果に関連付けられたタスクを抽出する手段と、
タスクの種類に関連付けられる複数のサブタスクを含むサブタスク知識を参照して、前記抽出されたタスクに関連付けられたサブタスク系列を求める手段と、
前記サブタスク系列の先頭のサブタスクを当該サブタスクとし、当該サブタスクの属するドメインを当該ドメインとして更新する手段と、
前記複数のドメインのそれぞれに含まれる行動知識のうち当該ドメインの行動知識を参照して、前記音声理解結果および前記サブタスクに関連付けられた行動またはサブタスク終了フラグを抽出する手段と、
前記抽出された行動を前記機械に実行させる手段と、
を備える装置。
- [2] 前記サブタスク知識は、前記タスクに関連付けられる1つ以上のサブタスクに関する知識と、前記サブタスクに関連付けられるドメインに関する知識と、を含む、請求項1に記載の装置。
- [3] 前記音声理解手段のそれぞれは、対応するドメインに関連性の高い複数の文パターンを含む音声知識を参照して、前記音声と前記複数の文パターンのそれぞれとの適合度を算出し、適合度が最も高い文パターンを選択し、前記選択された文パターンおよび該文パターンの適合度を音声理解結果として出力する、請求項1に記載の装置。

- [4] 前記選択する手段は、前記適合度に前記複数のドメインごとに設定された重みを乗じて信頼度を算出し、最大の信頼度をとるドメインを当該ドメインとして選択する、請求項3に記載の装置。
- [5] 前記更新する手段は、前記行動またはサブタスク終了フラグを抽出する手段がサブタスク終了フラグを抽出すると、前記サブタスク系列における当該サブタスクの次のサブタスクを新たな当該サブタスクとして更新し、新たな当該サブタスクの属するドメインを当該ドメインとして更新する、請求項1に記載の装置。
- [6] ユーザと機械とのインタラクションを実施するための方法であって、
マイクロフォンで検出された信号からユーザの音声を認識するステップと、
前記インタラクションにおける複数の段階に対応する複数のドメインのそれぞれに前記ユーザの音声を送るステップと、
前記複数のドメインのそれぞれにおいて、前記音声の内容を理解して音声理解結果を出力するステップと、
前記複数のドメインのそれぞれから前記音声理解結果を受け取るステップと、
前記複数の音声理解結果のうち最良の音声理解結果をとるドメインを、当該ドメインとして選択するステップと、
前記複数のドメインのそれぞれに含まれるタスク知識のうち前記当該ドメインのタスク知識を参照して、前記音声理解結果に関連付けられたタスクを抽出するステップと、
、
タスクの種類に関連付けられた複数のサブタスクを含むサブタスク知識を参照して、前記抽出されたタスクに関連付けられたサブタスク系列を求めるステップと、
前記サブタスク系列の先頭のサブタスクを当該サブタスクとし、当該サブタスクの属するドメインを当該ドメインとして更新するステップと、
前記複数のドメインのそれぞれに含まれる行動知識のうち当該ドメインの行動知識を参照して、前記音声理解結果および前記サブタスクに関連付けられた行動またはサブタスク終了フラグを抽出するステップと、
前記抽出された行動を前記機械に実行させるステップと、
を含む方法。

- [7] 前記サブタスク知識は、前記タスクに関連付けられる1つ以上のサブタスクに関する知識と、前記サブタスクに関連付けられるドメインに関する知識と、を含む、請求項6に記載の方法。
- [8] 前記出力するステップは、
対応するドメインに関連性の高い複数の文パターンを含む音声知識を参照して、前記音声と前記複数の文パターンのそれぞれとの適合度を算出するステップと、
前記適合度が最も高い文パターンを選択し、選択された文パターンおよび該文パターンの適合度を音声理解結果として出力するステップと、
を含む、請求項6に記載の方法。
- [9] 前記選択するステップは、前記適合度に前記複数のドメインごとに設定された重みを乗じて信頼度を算出し、最大の信頼度をとるドメインを当該ドメインとして選択するステップを含む、請求項8に記載の方法。
- [10] 前記更新するステップは、前記行動またはサブタスク終了フラグを抽出するステップにおいてサブタスク終了フラグが抽出されると、前記サブタスク系列における当該サブタスクの次のサブタスクを新たな当該サブタスクとして更新し、新たな当該サブタスクの属するドメインを当該ドメインとして更新するステップを含む、請求項6に記載の方法。
- [11] ユーザと機械とのインタラクションを実施するためのプログラムであって、
マイクロフォンで検出された信号からユーザの音声を認識する機能と、
ユーザとのインタラクションにおける複数の段階に対応する複数のドメインのそれぞれに前記ユーザの音声を送る機能と、
前記複数のドメインのそれぞれにおいて、前記音声の内容を理解して音声理解結果を出力する機能と、
前記複数のドメインのそれぞれから、前記音声理解結果を受け取る機能と、
前記複数の音声理解結果のうち最良の音声理解結果をとるドメインを、当該ドメインとして選択する機能と、
前記複数のドメインのそれぞれに含まれるタスク知識のうち前記当該ドメインのタスク知識を参照して、前記音声理解結果に関連付けられたタスクを抽出する機能と、

タスクの種類に関連付けられる複数のサブタスクを含むサブタスク知識を参照して、前記抽出されたタスクに関連付けられたサブタスク系列を求める機能と、

前記サブタスク系列の先頭のサブタスクを当該サブタスクとし、当該サブタスクの属するドメインを当該ドメインとして更新する機能と、

前記複数のドメインのそれぞれに含まれる行動知識のうち当該ドメインの行動知識を参照して、前記音声理解結果および前記サブタスクに関連付けられた行動またはサブタスク終了フラグを抽出する機能と、

前記抽出された行動を前記機械に実行させる機能と、
をコンピュータに実現させる、コンピュータ読取り可能な記録媒体に記録されたプログラム。

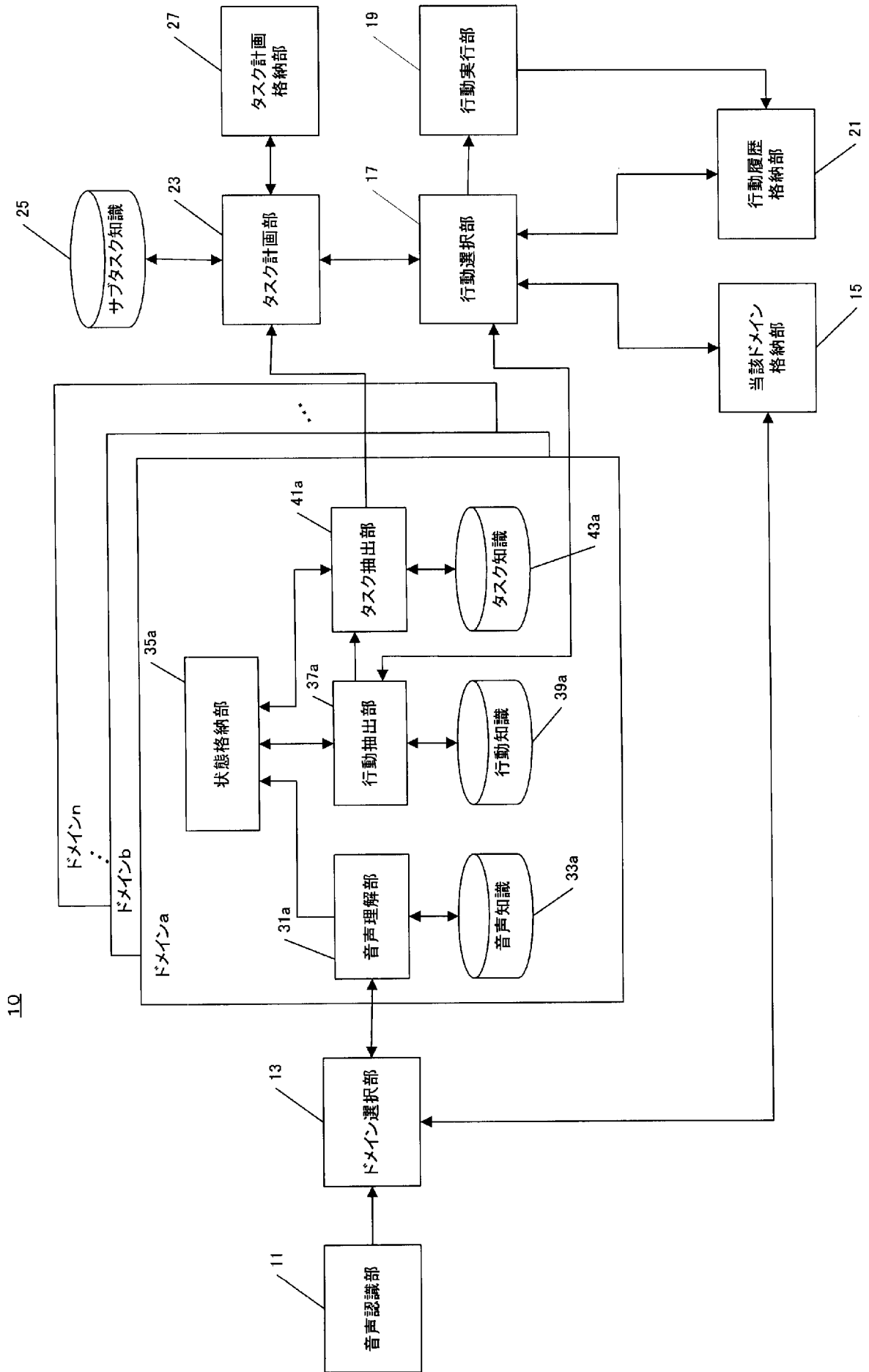
[12] 前記サブタスク知識は、前記タスクに関連付けられる1つ以上のサブタスクに関する知識と、前記サブタスクに関連付けられるドメインに関する知識と、を含む、請求項11に記載のプログラム。

[13] 前記出力する機能は、
対応するドメインに関連性の高い複数の文パターンを含む音声知識を参照して、前記音声と前記複数の文パターンのそれぞれとの適合度を算出する機能と、
前記適合度が最も高い文パターンを選択し、選択された文パターンおよび該文パターンの適合度を音声理解結果として出力する機能と、
を含む、請求項11に記載のプログラム。

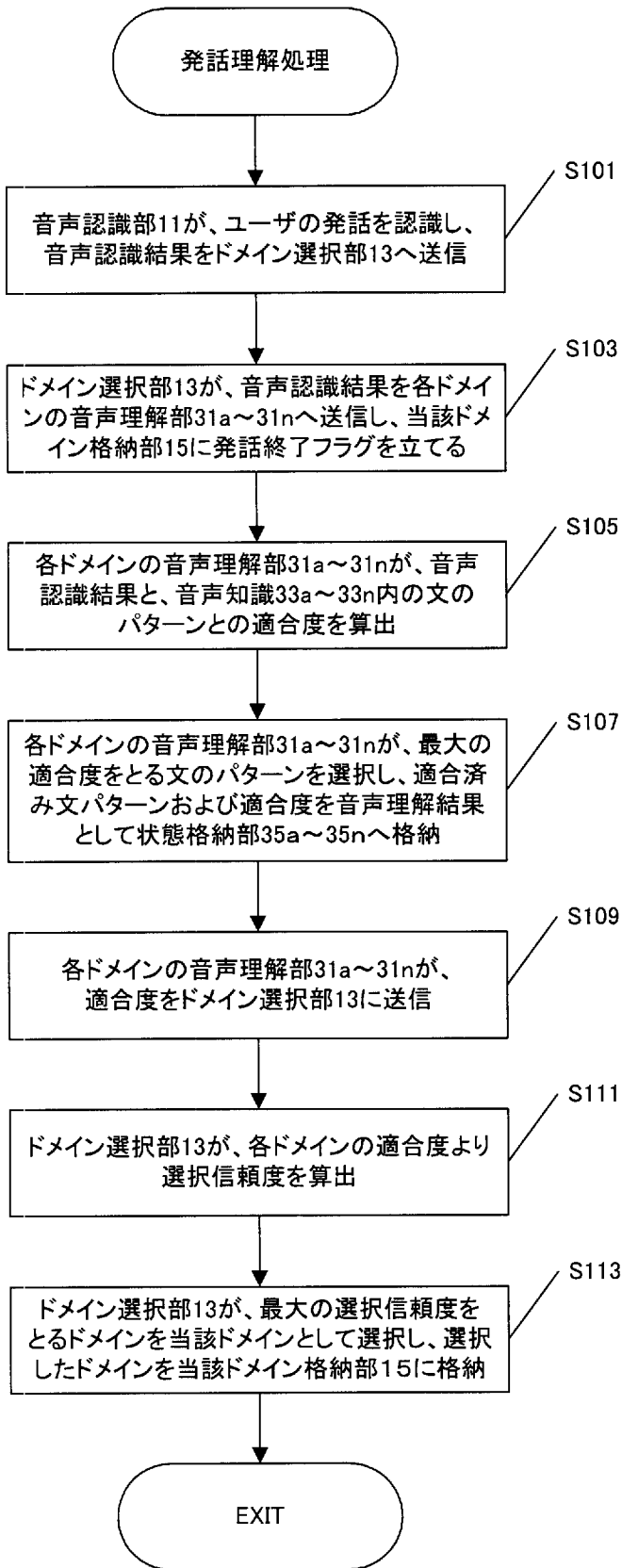
[14] 前記選択する機能は、前記適合度に前記複数のドメインごとに設定された重みを乗じて信頼度を算出し、最大の信頼度をとるドメインを当該ドメインとして選択する機能を含む、請求項13に記載のプログラム。

[15] 前記更新する機能は、前記行動またはサブタスク終了フラグを抽出する機能においてサブタスク終了フラグが抽出されると、前記サブタスク系列における当該サブタスクの次のサブタスクを新たな当該サブタスクとして更新し、新たな当該サブタスクの属するドメインを当該ドメインとして更新する機能を含む、請求項11に記載のプログラム。
。

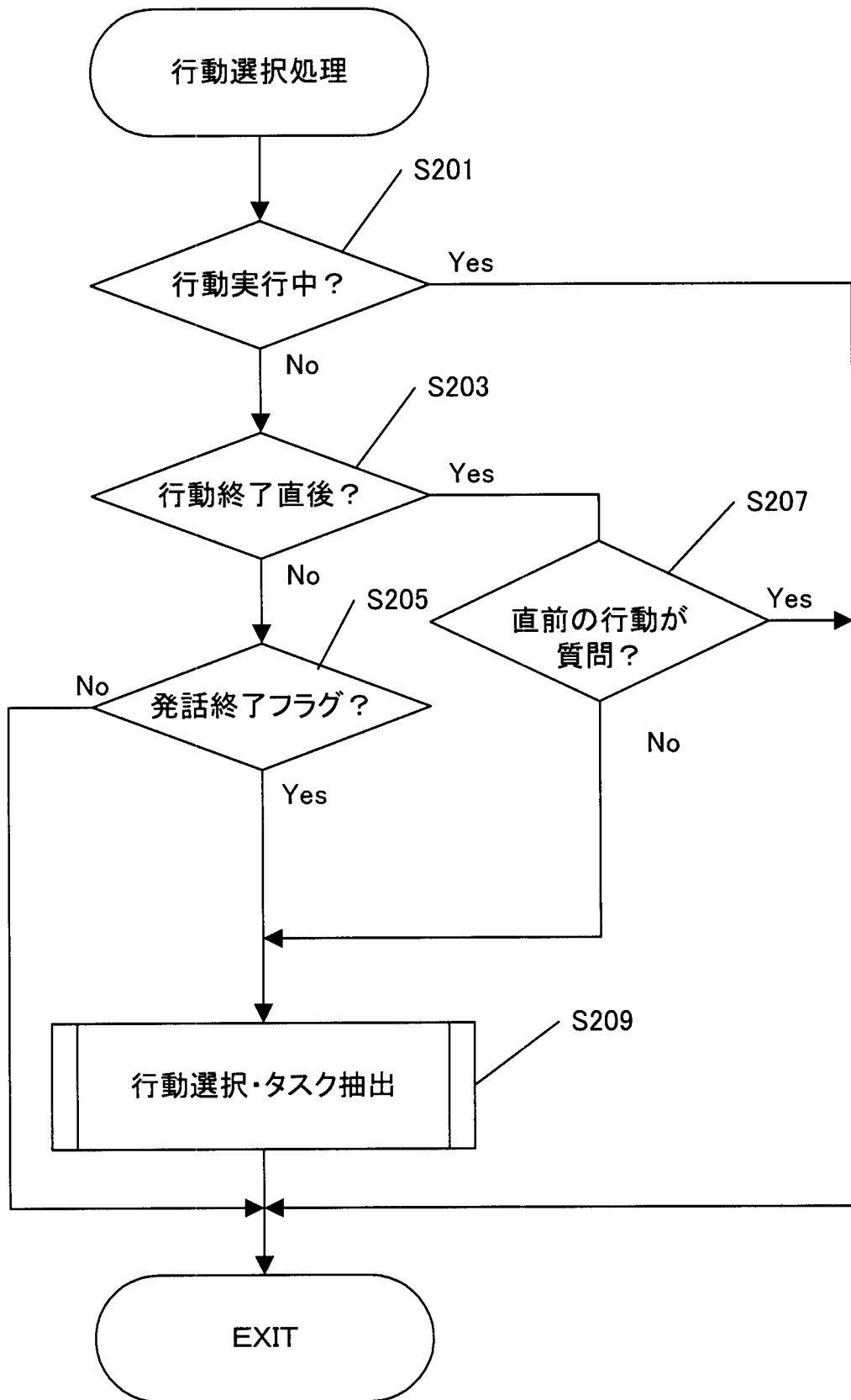
[図1]



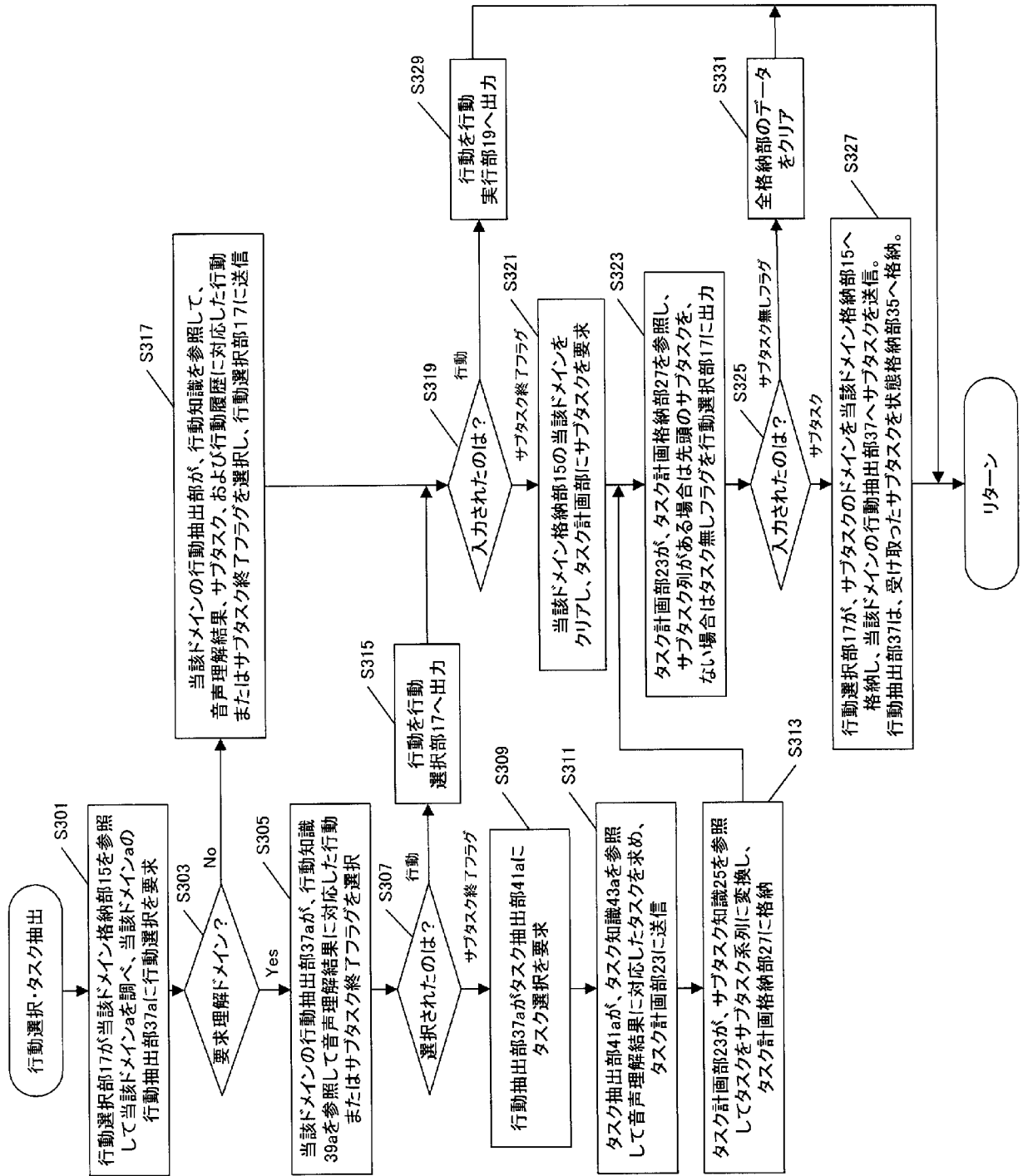
[図2]



[図3]



[図4]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2006/314786

A. CLASSIFICATION OF SUBJECT MATTER

G10L15/22(2006.01) i, G06F17/28(2006.01) i

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)

G10L15/22, G06F17/28

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

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2006
Kokai Jitsuyo Shinan Koho	1971-2006	Toroku Jitsuyo Shinan Koho	1994-2006

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

IEEE, JSTPlus (JDream2)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2004-526196 A (Voice-Insight), 26 August, 2004 (26.08.04), Full text; all drawings & WO 02/071393 A1 & US 2004/0064323 A1	1-15
A	JP 2003-316385 A (Hitachi, Ltd.), 07 November, 2003 (07.11.03), Full text; all drawings (Family: none)	1-15
A	JP 2001-343993 A (ATR Spoken Language Translation Research Laboratories), 14 December, 2001 (14.12.01), Full text; all drawings (Family: none)	1-15

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
10 August, 2006 (10.08.06)

Date of mailing of the international search report
22 August, 2006 (22.08.06)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2006/314786

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2001-356797 A (Nippon Telegraph And Telephone Corp.), 26 December, 2001 (26.12.01), Full text; all drawings (Family: none)	1-15
A	JP 2001-117921 A (Sony Corp.), 27 April, 2001 (27.04.01), Full text; all drawings (Family: none)	1-15
A	JP 2005-164836 A (Kabushiki Kaisha Kokusai Denki Tsushin Kiso Gijutsu Kenkyusho), 23 June, 2005 (23.06.05), Full text; all drawings (Family: none)	1-15
A	NAGAMORI et al., 'Multi Domain Onsei Taiwa System no Kochiku Shuho', Information Processing Society of Japan Kenkyu Hokoku [Spoken Language Processing], 02 June, 2000 (02.06.00), Vol.2000, No.54, 2000-SLP-31-7, pages 45 to 51	1-15
A	YASUDA et al., 'Tan'itsu Domain System no Togo ni yoru Fukusu Domain Onsei Taiwa System', Information Processing Society of Japan Kenkyu Hokoku [Spoken Language Processing], 07 February, 2003 (07.02.03), Vol.2003, No.14, 2003-SLP-45-20, pages 119 to 124	1-15
A	KAWAGUCHI et al., 'Fukusu no Onsei Taiwa System no Togo Seigyo Kiko to Sono Hyoka', Information Processing Society of Japan Kenkyu Hokoku [Spoken Language Processing], 01 June, 2001 (01.06.01), Vol.2001, No.55, 2001-SLP-36-10, pages 63 to 70	1-15
T	KANDA et al., 'Fukusu Domain Onsei Taiwa System ni Okeru Taiwa Rireki o Riyo shita Domain Sentaku no Koseidoka', Information Processing Society of Japan Kenkyu Hokoku [Spoken Language Processing], 03 February, 2006 (03.02.06), Vol.2006, No.12, 2006-SLP-60(11), pages 55 to 60	1-15

A. 発明の属する分野の分類 (国際特許分類 (IPC)) Int.Cl. G10L15/22(2006.01)i, G06F17/28(2006.01)i		
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)) Int.Cl. G10L15/22, G06F17/28		
最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1922-1996年 日本国公開実用新案公報 1971-2006年 日本国実用新案登録公報 1996-2006年 日本国登録実用新案公報 1994-2006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) IEEE, JSTPlus(JDream2)		
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 請求の範囲の番号
A	J P 2 0 0 4 - 5 2 6 1 9 6 A (ヴォイス-インサイト) 2004.08.26, 全文, 全図 & WO 02/071393 A1 & US 2004/0064323 A1	1-15
A	J P 2 0 0 3 - 3 1 6 3 8 5 A (株式会社日立製作所) 2003.11.07, 全文, 全図 (ファミリーなし)	1-15
<input checked="" type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> パテントファミリーに関する別紙を参照。		
* 引用文献のカテゴリー 「A」特に関連のある文献ではなく、一般的技術水準を示すもの 「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの 「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す) 「O」口頭による開示、使用、展示等に言及する文献 「P」国際出願日前で、かつ優先権の主張の基礎となる出願日の後に公表された文献 「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの 「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの 「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの 「&」同一パテントファミリー文献		
国際調査を完了した日 10.08.2006		国際調査報告の発送日 22.08.2006
国際調査機関の名称及びあて先 日本国特許庁 (ISA/J P) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号		特許庁審査官 (権限のある職員) 榎本 剛 電話番号 03-3581-1101 内線 3541
		5 Z 9 3 7 9

C (続き) . 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 請求の範囲の番号
A	J P 2001-343993 A (株式会社エイ・ティ・アール 音声言語通信研究所) 2001.12.14, 全文, 全図 (ファミリーなし)	1-15
A	J P 2001-356797 A (日本電信電話株式会社) 2001.12.26, 全文, 全図 (ファミリーなし)	1-15
A	J P 2001-117921 A (ソニー株式会社) 2001.04.27, 全文, 全図 (ファミリーなし)	1-15
A	J P 2005-164836 A (株式会社国際電気通信基礎技 術研究所) 2005.06.23, 全文, 全図 (ファミリーなし)	1-15
A	長森他、「マルチドメイン音声対話システムの構築手法」、情報処理 学会研究報告 [音声言語情報処理], 2000.06.02, Vol.2000, No.54, 2000-SLP-31-7, p.45-51	1-15
A	安田他、「単一ドメインシステムの統合による複数ドメイン音声対話 システム」、情報処理学会研究報告 [音声言語情報処理], 2003.02.07, Vol.2003, No.14, 2003-SLP-45-20, p.119-124	1-15
A	河口他、「複数の音声対話システムの統合制御機構とその評価」、情 報処理学会研究報告 [音声言語情報処理], 2001.06.01, Vol.2001, No.55, 2001-SLP-36-10, p.63-70	1-15
T	神田他、「複数ドメイン音声対話システムにおける対話履歴を利用し たドメイン選択の高精度化」、情報処理学会研究報告 [音声言語情報 処理], 2006.02.03, Vol.2006, No.12, 2006-SLP-60(11), p.55-60	1-15